REF
QK1
.M833

FLORISTICS FOR THE 21st CENTURY

Proceedings of the workshop
4–7 May, 1988
Alexandria, Virginia

Edited by Nancy R. Morin,
R. David Whetstone, Dieter Wilken,
and Kay L. Tomlinson

Missouri Botanical Garden

Cover art by John Myers, Missouri Botanical Garden, based on
Plate 154 of *Floerkea proserpinacoides* Willd. in
Asa Gray, *The Genera of the Plants of the United States.*
New York: George P. Putnam. 1849.

FLORISTICS FOR THE 21st CENTURY

FLORISTICS FOR THE 21st CENTURY

Proceedings of the workshop sponsored by the
American Society of Plant Taxonomists and the
Flora of North America Project
4–7 May, 1988
Alexandria, Virginia

Edited by
Nancy R. Morin, R. David Whetstone, Dieter Wilken, and Kay L. Tomlinson

Monographs in Systematic Botany from the Missouri Botanical Garden

MISSOURI BOTANICAL

SEP 7 1989

GARDEN LIBRARY

MONOGRAPHS IN SYSTEMATIC BOTANY
FROM THE MISSOURI BOTANICAL GARDEN

Volume 28

June 1989

ISSN 0161-1542

Copyright © 1989 by Missouri Botanical Garden
All rights reserved.

The workshop "Floristics for the 21st Century" and the production of this report were supported by grants from the National Science Foundation, the David and Lucile Packard Foundation, The Pew Charitable Trusts, and by the Missouri Botanical Garden.

Printed in the United States of America by
Braun-Brumfield, Inc., Ann Arbor, Michigan

TABLE OF CONTENTS

Preface vii

List of Participants x

Workshop Goals 1

Introduction 3

Concepts 5

Part I: Presentations

Flora Europaea and the European Documentation System 8

The PRECIS Computer System and its Applications for the Flora of Southern Africa 11

Flora of Australia 23

Other Floristic and Related Projects of Various Scopes 36

Distribution Information for Management 42

Flora of North America Project 44

Part II: Synopsis of Group Recommendations

How Floristic Information is Used 58

Considerations for Producing the Ideal Flora 71

Conclusions: New Directions for Floristic Research 90

Appendix 1. Report of Group 1: Agriculture, Forestry, and Horticulture 94

Appendix 2. Report of Group 2: Assessing Environmental Conditions 100

Appendix 3. Report of Group 3: Assessing Responses to Changes in Environment 103

Appendix 4. Report of Group 4: Biological Sciences Other than Plant Systematics 112

Appendix 5. Report of Group 5: Phylogeny and Biogeography 122
Appendix 6. Report of Group 6: Resources 129
Appendix 7. Report of Group 7: Computer Resources 135
Appendix 8. Report of Group 8: Interaction and Collaboration 143
Appendix 9. Report of Group 9: Long-term Management 148
Appendix 10. Report of Group 10: Additional Topics 153
Appendix 11. Workshop Goals and Agenda 158

PREFACE

Botanists have been writing Floras, which are publications containing scientific accounts of the plants in defined areas, since the early 1600s. Physicians used Floras for information on medicinal plants and, later, horticulturalists used Floras to identify plants that might be brought into cultivation. Floras have documented plants collected in exotic places as well as plants growing in a botanist's "backyard." The form of Floras, lists of plants with associated information about them, has changed very little over the centuries, but in the last two decades the use of computers to store and organize such information has increased dramatically.

Early Floras usually were written by single individuals based on observations of the limited plant samples available and study of existing, also limited, literature. More recently, Floras of large scope have been written by collaboration of many authors who collectively have had to examine thousands of plant samples and to evaluate and incorporate information from dozens, or even hundreds, of publications. Interest in tropical areas and concern about the environment have stimulated increased efforts to inventory and document the plants of the world. The growing realization that we must both conserve our rapidly dwindling resources and use them more wisely gives an added urgency to these efforts. Thus, the amount of time and money needed for floristic work is increasing but so are the diversity and importance of ways in which floristic information is being used and the number of people who have a direct interest in floristic data.

One of the major, collaborative floristic efforts underway is the Flora of North America project. Its intent is to synthesize existing information about the higher plants (that is, excluding mosses, liverworts, algae, and fungi) of North America north of Mexico. The project will produce both published volumes and a computerized database and is expected to take about 12 years to complete. The amount of effort that will be required collectively is enormous, but the potential usefulness of the project, particularly in view of the use of computers, is almost unlimited.

Because of the long-term commitment of funds and manpower that will be required for this project and that is already being devoted to other floristic projects, and because of the new opportunities presented by computer-assisted methods, the National Science Foundation and the American Society of Plant Taxonomists considered that this was an important time to take stock of where we are in terms of providing and using floristic information, and where we should be heading. They suggested that the Flora of North America project organize a workshop, the participants in which would be professionals from a very wide range of existing or potential user groups, in order to determine how floristic information is being used now and what could be done to make the information more useful and to encourage its use.

"Floristics for the 21st Century" was held in Alexandria, Virginia, May 4--7, 1988. The workshop was co-sponsored by the American Society of Plant Taxonomists and the Flora of North America project, and it was hosted by the American Institute of Biological Sciences. It was funded by the National Science Foundation; participation

by Flora of North America project staff was supported by the David and Lucile Packard Foundation. About 70 participants represented fields from agriculture to zoology, including both applied and basic research (Appendix 11). Although the focus of the discussion was the Flora of North America project, participants realized that the workshop conclusions could be extended to organisms other than plants and to areas other than North America.

The workshop was designed to inform participants about current projects and especially about the uses that are being made of computerized databases, to help them evaluate their current use of information about plants, and to bring out their creative ideas about how such information could be presented best to suit their needs and to increase the use they could make of it. The full schedule is found in Appendix 11. On the first day participants were introduced to the Flora of North America project and to other floristic and biological inventory projects now underway. They met in groups on the second day to discuss how they used information about plants currently in their jobs and to brainstorm about what the ideal content and presentation of the information might be. At the end of the day, the group met as a whole, each working group presenting its ideas and conclusions to the rest. On the third day new working groups were formed to discuss ways in which the needs identified the day before could be filled: what workforce would be required, how computers could be used, how projects might collaborate, and what the long-term management needs would be. At the end of the day the entire group met again to compare notes. The last day was spent writing group reports and summarizing the findings.

This document contains some of the background material presented at the workshop, a synthesis of the ideas and conclusions brought out by the working groups, a general summary, and the working group reports. Presentation reports were provided by some authors; the rest were prepared from notes taken by John Packer and Robert Kiger. The synopsis was drawn from the working group reports and from original materials from the working groups. General ideas were pulled together and summarized by Dieter Wilken, David Whetstone, and myself, and then Kay Tomlinson and I rewrote and edited as necessary. The working group reports were written initially at the workshop, then circulated among the group members, edited by the group leader and recorder, and then submitted for final editing. Although an effort has been made at every stage to maintain the original ideas in each report, wording and style may have been modified somewhat in the editing process.

Acknowledgments: Louise Salmon of the American Institute of Biological Sciences and Kay Tomlinson of the Flora of North America project handled all arrangements. The success of the workshop is in large part due to their efforts. Patricia Rich Associates helped design the workshop and trained the leaders and recorders. This professional expertise helped greatly to make the workshop effective. Flora of North America editorial committee members (George Argus, Theodore Barkley, Luc Brouillet, James Estes, John Fay, Ronald Hartman, Marshall Johnston, Robert Kiger, David Murray, John Packer, Leila Shultz, Richard Spellenberg, and David Whetstone) provided information and help throughout the workshop. Group leaders and recorders played an especially important role in leading the discussions and brainstorming sessions and in recording, verifying, and writing the groups' ideas and conclusions. Leaders and recorders were: George Argus, Theodore Barkley, Luc Brouillet, James Estes, John Fay, Ronald Hartman, Marshall Johnston, Robert Kiger, Michael Macrander, David Murray, Ann Rhoads, Jennifer Richards, Leila Shultz, Beryl Simpson, Richard Spellenberg, Tod Stuessy, Gary Waggoner, David Whetstone,

Dieter Wilken, and David Young. Marshall Crosby and Robert Magill, both of the Missouri Botanical Garden, helped throughout the workshop, particularly with the computer demonstrations. Several participants travelled considerable distances to attend the workshop: Alex George (Australia), G. E. Gibbs Russell (South Africa), and Vernon Heywood (England). In the end it was the work of all the participants that made the workshop an outstanding success, and the American Society of Plant Taxonomists and the Flora of North America project are deeply grateful to each and every one of them for the time and energy that they devoted to it. We are also very grateful for the financial support provided by the National Science Foundation, the David and Lucile Packard Foundation, and The Pew Charitable Trusts.

Nancy R. Morin
St. Louis, Missouri
11 January 1989

FLORISTICS FOR THE 21st CENTURY

LIST OF PARTICIPANTS

Dr. George W. Argus
National Herbarium
National Museums of Canada
Ottawa, Ontario
Canada K1P 6P4

Dr. Theodore M. Barkley
Division of Biology
Ackert Hall
Kansas State University
Manhattan, Kansas 66506

Dr. Chris Beecher
Napralert
College of Pharmacy
833 South Wood Street
Chicago, Illinois 60612

Dr. André Bouchard
Institut Botanique
Université de Montréal
4101 est, rue Sherbrooke
Montréal, Québec
Canada H1X 2B2

Dr. Warren U. Brigham
State Natural History Survey Division
Natural Resources Building
607 East Peabody Drive
Champaign, Illinois 61820

Dr. Luc Brouillet, Conservateur
Herbier Marie-Victorin
Institut Botanique
Université de Montréal
Montréal, Québec
Canada H1X 2B2

Dr. Paul M. Catling
Central Experimental Farm
Agriculture Canada
Ottawa, Ontario
Canada K1A 0C6

Dr. Barry Chernoff
Fishes, Department of Zoology
Field Museum of Natural History
Roosevelt Road at Lake Shore Drive
Chicago, Illinois 60605

Dr. Marshall R. Crosby
Missouri Botanical Garden
P. O. Box 299
St. Louis, Missouri 63166

Dr. Thomas Duncan
Department of Botany
University Herbarium
University of California
Berkeley, California 94720

Mr. Donald N. Duvick
Senior Vice President of Research
Pioneer Hi-Bred International, Inc.
P. O. Box 81
Johnston, Iowa 50131

Dr. James R. Estes
Bebb Herbarium
University of Oklahoma
770 Van Vleet Oval
Norman, Oklahoma 73019

Dr. John Fay
Office of Endangered Species
U.S. Fish and Wildlife Service
Washington, D.C. 20240

Dr. Harrison L. Flint
Department of Horticulture
Purdue University
West Lafayette, Indiana 47907

Mr. Alexander S. George
Executive Editor
Flora of Australia
G.P.O. Box 1383
Canberra, A.C.T. 2601
Australia

Dr. G. E. Gibbs Russell
Botanical Research Institute
Private Bag X101
Pretoria 0001
South Africa

Dr. Arthur C. Gibson
Department of Biology
University of California
Los Angeles, California 90024

Dr. Robert H. Giles, Jr.
Department of Forestry
Virginia Polytechnic Institute and
State University
Blacksburg, Virginia 24061

Dr. Robin Lambert Graham
Environmental Science Division
Oak Ridge National Laboratory
P. O. Box X
Oak Ridge, Tennessee 37830

Dr. Charles R. Gunn
Systematic Botany, Mycology, and
Nematology Laboratory-USDA/ARS
Building 265, BARC-East
Beltsville, Maryland 20705

Dr. David W. Hall
Department of Botany
University of Florida
209 Rolfs Hall
Gainesville, Florida 32611

Dr. Ronald L. Hartman
Rocky Mountain Herbarium
University of Wyoming
3165 University Station
Laramie, Wyoming 82071

Dr. Don W. Hayne
Institute of Statistics
North Carolina State University
P. O. Box 8203
Raleigh, North Carolina 27695-8203

Dr. Robert R. Haynes
Department of Biology
University of Alabama
Tuscaloosa, Alabama 35486-1927

Prof. Vernon H. Heywood
IUCN Conservation Monitoring
Centre
53 The Green
Kew, Richmond, Surrey TW9 3AA
United Kingdom

Dr. James C. Hickman
Jepson Herbarium
Department of Botany
University of California
Berkeley, California 94720

Dr. Marshall C. Johnston
Plant Resources Center
The University of Texas at Austin
Austin, Texas 78713

Dr. John Kartesz
North Carolina Botanical Garden
University of North Carolina
Chapel Hill, North Carolina 27514

Dr. Robert W. Kiger
Hunt Institute for Botanical
Documentation
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

Dr. Michael Kosztarab
Department of Entomology
Virginia Polytechnic Institute
and State University
Blacksburg, Virginia 24061

Dr. Stanley L. Krugman
U.S. Forest Service
P. O. Box 2417
Washington, D.C. 20013

Dr. Gary Larson
Department of Biology
South Dakota State University
Box 2207B
Brookings, South Dakota 57007-0595

Dr. Andrew Leiser
Department of Ornamental Horticulture
University of California
Davis, California 95616

Dr. Bruce MacBryde
U.S. Fish & Wildlife Service
Office of Scientific Authority
Washington, D.C. 20240

Dr. Michael Macrander
Department of Biology
University of Alabama
Tuscaloosa, Alabama 35486-6282

Dr. Robert E. Magill
Missouri Botanical Garden
P. O. Box 299
St. Louis, Missouri 63166

Dr. Norton G. Miller
Biological Survey
New York State Museum & Science
Service
Cultural Education Center
Albany, New York 12230

Dr. Stanley A. Morain
Technology Application Center
University of New Mexico
Albuquerque, New Mexico 87131

Dr. Nancy R. Morin
Missouri Botanical Garden
P. O. Box 299
St. Louis, Missouri 63166

Dr. Larry E. Morse
The Nature Conservancy
1800 North Kent Street
Arlington, Virginia 22209

Dr. David F. Murray
Herbarium
University of Alaska Museum
Fairbanks, Alaska 99701

Dr. Curtis J. Nelson
Department of Agronomy
University of Missouri
Columbia, Missouri 65211

Dr. Lorin I. Nevling
Illinois Natural History Survey
607 East Peabody Drive
Champaign, Illinois 61820

Dr. John G. Packer
Department of Botany
University of Alberta
Edmonton, Alberta
Canada T6G 2E9

Dr. Ann F. Rhoads
Morris Arboretum
University of Pennsylvania
9414 Meadowbrook Avenue
Philadelphia, Pennsylvania 19118

Dr. Jennifer H. Richards
Department of Biological Sciences
Florida International University
Miami, Florida 33199

Dr. Stanwyn G. Shetler
Department of Botany, NHB 166
Smithsonian Institution
Washington, D.C. 20560

Dr. Leila M. Shultz
Department of Biology
Utah State University
UMC 55
Logan, Utah 84322

Dr. Beryl B. Simpson
Department of Botany
University of Texas
Austin, Texas 78713

Dr. Richard W. Spellenberg
Department of Biology
New Mexico State University
Las Cruces, New Mexico 88003-0032

Dr. Bruce Stein
The Nature Conservancy
1815 North Lynn Street
Arlington, Virginia 22209

Dr. Tod F. Stuessy
Department of Biology
Ohio State University
1735 Neil Avenue
Columbus, Ohio 43210-1293

Dr. Iain E. P. Taylor
Department of Botany
University of British Columbia
Vancouver, British Columbia
Canada V6T 1W5

Dr. Kay L. Tomlinson
Missouri Botanical Garden
P. O. Box 299
St. Louis, Missouri 63166

Dr. Gary Waggoner
Geographic Information
Systems Division
National Park Service
Denver, Colorado 80225

Dr. Jess Wagstaff
Food & Drug Administration
HFF-108
200 C Street, SW
Washington, D.C. 20204

Dr. Kerry S. Walter
Center for Plant Conservation
125 The Arborway
Jamaica Plain, Massachusetts 02130

Dr. R. David Whetstone
Department of Biology
Jacksonville State University
Jacksonville, Alabama 36265

Dr. Dieter Wilken
Department of Biology
Colorado State University
Fort Collins, Colorado 80523

Dr. David A. Young
Santa Barbara Botanic Garden
1212 Mission Canyon Road
Santa Barbara, California 93105

National Science Foundation
Observers:

Dr. Robert J. Robbins
Division of Biotic Systems and
Resources
National Science Foundation
1800 G Street, NW
Washington, D.C. 20550

Dr. James E. Rodman
Program Director
Systematic Biology Program
National Science Foundation
1800 G Street, NW
Washington, D.C. 20550

Dr. David E. Schindel
Program Director
Systematic Biology Program
National Science Foundation
1800 G Street, NW
Washington, D.C. 20550

FLORISTICS FOR THE 21st CENTURY

WORKSHOP GOALS

-- **James E. Rodman**

The current Flora of North America project is clearly a major enterprise, by the community, by the National Science Foundation (NSF), and, I would hope, by other funding sources. In that light, it is particularly encouraging that support has been received from private sources such as the J. N. Pew, Jr., Charitable Trust and the David and Lucile Packard Foundation. Flora of North America builds upon a long history of vision, persuasion, frustration, defeats, small triumphs, and perseverance. All of that is likely to continue, with--we all hope--the small triumphs translating into larger achievements. Workshop participants will help in this endeavor by defining the dimensions of the task for a modern, continental floristic inventory--what kinds of information do we want and need, how do we acquire the information, who uses it and how, and who will use it in the future.

The current Flora of North America enterprise is envisaged to take 10 to 12 years, a very optimistic timetable, and one critically dependent upon a great deal of volunteer labor. The Systematic Biology Program at NSF, on its own, cannot guarantee funding over that long time period, nor will its likely budgets pay for anything beyond a bare-bones endeavor. It is no secret that this workshop report must serve as a strong argument for increased support for a Flora of North America project, with computer databases as a major focus. Whether this increased support is provided by the Systematic Biology Program at NSF or by other agencies or funding sources is secondary, it seems to me, to the need to articulate strong arguments for the value and utility of floristic information in formats that make it fast and easy to use. With those strong arguments in hand, systematists and their colleagues can begin to convince funding sources to provide additional and enduring support.

The National Science Foundation would like this report to contain an appraisal: first, what *is* floristic information, who produces it, and who uses it, and how do they use it? Many of us grew up on Fernald's Manual or Jepson's, or the Canadian analog, or the Australian one; but those are just parts of a larger picture. What is that larger picture of floristic information, for North America north of Mexico? Second, and more challenging, what *could* floristic information be, who will produce that new product, who will the users be, and how will they use it?

Obviously, computers loom large in this endeavor, or at least they should. But the task of this workshop is *not* to design computer databases of floristic information. That will come in another round of discussions and workshops, building upon this one. The task of this workshop is to set the parameters for the work that is to be done. The report of this workshop will be the guiding document for the eventual design and implementation of major computer databases of floristic information, with emphasis on the plural. Those databases will include both taxon- and specimen-based systems, with eventual connections to Geographic Information Systems, and they will likely

require more than one central location, with networking among large and small institutions a significant feature.

Currently, the Flora of North America machinery is in place to produce a rather traditional floristic product, and to build computer resources to aid that task. Looking five to ten years ahead, we can all predict that the present distinction between published floristic information and data in computer systems will blur and then disappear. Rather, dissemination from computer databases will be fast and easy and will take many forms, some of which we can anticipate now, for instance compact discs, networked files, and, yes, books, and other formats yet to be invented.

Knowing what floristics researchers do now, what constituencies they serve and how, we can go on to suggest improvement and expansion of those tasks, and keep the enterprise open to future demands. That is the charge of this ambitiously titled workshop and its report. And I have every confidence it will be met.

INTRODUCTION

-- Tod F. Stuessy

The American Society of Plant Taxonomists (ASPT) has been closely involved with efforts to write a Flora of North America for several decades. At times the relationship between ASPT and groups endeavoring to organize this project has seemed like a very delicate marriage, in which at one moment everything has appeared rosy, and at other moments on the verge of divorce. A certain amount of intrafamilial squabbling is to be expected in any enterprise, but the important perspective is that the relationship is now taking a more romantic tone. As evidence that the ASPT is solidly behind the Flora of North America project, let me point out that it is our membership who will be doing most of the work! The professional taxonomists of this country and Canada are those who will be participating in writing various parts of the overall project. Further, all of those on the editorial committee are members of the ASPT. We also are enthusiastic co-sponsors of the workshop that has produced this excellent report. And finally, our Environmental and Public Policy Committee, this year chaired by Dieter Wilken, has just produced an inventory of ASPT members who have expertise in rare and endangered species as well as broader conservation issues.

There are many reasons why the ASPT is now solidly behind the Flora of North America project. A project of this scope allows our society to be more responsive to the needs of people in this country and Canada in general from the standpoint of improved land management, conservation, weed control, and so on. In the long term, this will benefit our overall exposure and access to research funds, which will enable us to increase our research productivity and enhance our longer-term benefit to society. In the short run, it is likely that some additional jobs will be available for professional taxonomists, which is one of the most important concerns within our membership. Flora of North America should also stimulate floristic inventory at the state and local levels, and perhaps will even stimulate the development of the National Biological Survey, which is now receiving attention. Flora of North America should also develop a credible base of information that will allow us to ask more focused and more meaningful biological questions about evolutionary aspects of our flora. It is difficult to analyze biogeographic and evolutionary questions with precision when the database on relationships, geography, and associated environmental variables is incomplete.

The ASPT is presently increasing its activities through a new Five-Year Plan of Development. Among our objectives is to obtain funds for: field research; herbarium visits; the publishing of books and symposia; the development of manuals of funding sources, teaching resources, and endangered species expertise; the holding of regional symposia; the development of a network of herbaria; and so on. Participating in Flora of North America is totally and completely compatible with our new emphasis on increased activities. The objective of this present workshop--to gain guidance from those who will use Flora of North America in order to improve its value--is completely compatible with the kind of responsiveness to society in general that the ASPT is now seeking.

The ASPT has great interest in "Floristics for the 21st Century." Our membership is involved with many floristic projects worldwide, and we view the Flora of North

America project as a chance to see new model programs developed that will have broader application to other efforts on a regional scale or in other areas of a smaller nature throughout the world. Databasing, new types of data, new presentations of data, and other topics are all of great concern to those people working on floristic projects within ASPT. We are, therefore, extremely interested in looking toward innovations that will come from the broader effort that will then be applicable to many other projects throughout our plant taxonomic community.

A word of caution, however, might profitably be stressed at this time. Whatever models for new approaches or database strategies emerge from Flora of North America, and however positive these will be, the main objective must be to produce the twelve volumes of hard copy so desperately needed by teachers, researchers, evolutionary biologists, conservationists, land managers, city planners, developers and land investors, and so forth. We have learned our lesson from the past two decades of efforts to write a Flora of North America, and I think we all know how important maintaining the hard copy objective is. Technology has changed in the meantime so that to proceed with Flora of North America at this time without using computers to develop searchable databases would be irresponsible--as would be starting still another round of creating databases and futuristic modeling without the production of the final useful volumes.

It is a real pleasure for me as the President of the ASPT to be able to introduce this workshop volume. Flora of North America is a project whose time has come. It needs the ASPT, and the ASPT needs the Flora of North America project. The overwhelming good news about funding from the National Science Foundation plus the additional funds that will be matched from the J. N. Pew, Jr., Charitable Trust insures that the project is on its way in good fashion. The future funding will be difficult and will require hard work and cooperation among members of our taxonomic community. The ASPT is solidly and squarely behind this most important project, and we intend to do everything possible to aid its development and successful completion.

CONCEPTS

-- Nancy R. Morin

The word "flora" refers to the plants occurring within a given region as well as to a publication describing those plants. In order to distinguish between the two, the word is sometimes capitalized when a publication is meant: "Flora," and this convention is used throughout this report. (Zoologists use a similar convention--fauna and Fauna.) A Flora may contain anything from a simple list of the plants occurring in an area to a very detailed account of those plants. Floras are different from popular manuals in that they attempt to cover all of the plants, rather than only the most common or conspicuous ones.

A Flora almost always contains scientific names, and it may also include vernacular (or "common") names, literature references, descriptions, habitats, geographical distribution, illustrations, flowering times, or notes. Less often, Floras include chemical information, chromosome numbers, population occurrences, etc. Sometimes the plants are listed alphabetically, and sometimes they are presented within a classification system that indicates which plants are most similar or are thought to be related. Floras often also include identification devices. These are usually "keys" that consist of pairs of mutually exclusive statements. The statement chosen leads the user either to more statements or to the identity of the plant.

Workshop participants discussed in some detail the meaning of the word "Flora" to be used for the purposes of the workshop. Some argued that the most traditional sense should be maintained, and that a Flora at most includes names, descriptions, habitat, and distribution. Others believed that the Flora concept could include an unlimited amount of additional information about the plants. Subsequent deliberations were based on this second, broader definition.

Aspects of Floras that were considered in the workshop include (1) nomenclature (what scientific names apply, or did apply, and the manner in which they were applied, to the included plants), (2) classification (to what other plants a particular kind of plant is thought to be related), (3) literature (basic published references), (4) descriptions (the primary attributes of the plants), (5) identification devices (e.g., keys, comparison charts, illustrations), (6) habitat (ecological conditions under which the plant grows), (7) phenology (times of flowering and fruiting), (8) distribution (where the plants grow), and (9) ethnobotany (aspects of special interest to man--medicinal or horticultural uses, toxicity, etc.). Some Floras are synoptical--they contain only information essential for identification and basic characterization of a plant. Others are revisionary or monographic in scope; that is, they include very detailed and extensive descriptions and distributions. Either kind of Flora may reflect new interpretations of the groups and their relationships based on field, library, herbarium, and laboratory investigations.

Floras are usually written by plant taxonomists, that is, botanists who study the classification of plants. Currently taxonomy incorporates the understanding that plants are related, that they are descended from ancestral plants, and that they have evolved over time. Through time and evolution the offspring of closely related plants have become less similar and have lost the ability to exchange genetic

information. What we see in nature are patterns of discontinuous variation and differing degrees of similarity and ability to interbreed.

In the hierarchical system used by botanists, closely similar plants that (usually) are still capable of interbreeding are grouped into species. Similar, or related, species are grouped into genera, and genera sharing basic characteristics are grouped into families. Any circumscribed group at any of these ranks (e.g., species, genus, family) is a taxon (plural: taxa). In most Floras the taxa are presented in some form of hierarchical classification, and within a rank they may be arranged either according to hypothesized relationships or alphabetically. Thus, families may be arranged alphabetically, within each one the included genera may be alphabetical, and within each genus the species may be alphabetical. Or, families may be arranged taxonomically but their included genera arranged alphabetically. Taxonomic arrangements contain more information than alphabetical ones because they convey how the groups are thought to be related and what characteristics those groups share. Scientific names themselves contain some of this information: the name of a species has three parts--the genus name (e.g., *Rosa*), the specific epithet (e.g., *carolina*), and the name of the person who described the species (e.g., Linnaeus). The full species name is *Rosa carolina* Linnaeus. The scientific name of the plant shows what genus it belongs in and implies that the plant shares general characteristics found in other members of that genus. To find additional information about a plant, or to communicate with others about that plant, it is *essential* that the correct scientific name be known and used. Thus, works such as Floras perform an immediate practical function: identifying plants and assigning their correct names. Information presented in a Flora may be for the taxon, for instance for the family or species as a whole, or for the individual plant. Taxon descriptions are general and usually include ranges of variation, e.g., leaves linear to oval, 2--4 millimeters long. Thus, some plants of the species might have linear leaves that are 3.5 millimeters long whereas others may have leaves that are nearly oval and 2 millimeters long. The shapes and dimensions are for the general case, and are not linked to specific individual specimens. The geographic distribution for the taxon is also generalized, for example "California, Oregon, and Washington"; or, possibly, "Pacific coastal areas."

Information on individual plants is much more precise. Although a range of measurements may still be given, the measurements are linked to an individual specimen--an actual object rather than a concept. Distributional information is similarly precise, for example "California, Sacramento County, Sacramento, 2.5 miles west of the intersection of U.S. Highway 80 and U. S. Highway 5 on Highway 80, north-facing roadbank, growing with *Brassica nigra* (Linnaeus) Koch and *Erodium cicutarium* (Linnaeus) L'Héritier ex Aiton.[1] Discussions of specimen databases found in this report refer primarily to this precise locality information, captured directly from herbarium specimens.[2] Databases will allow fairly high-resolution mapping of each species and subsequent analysis of distribution data.

---

[1] Although these names look complicated, they contain a great deal of information. *Brassica nigra* is the common yellow-flowered mustard introduced from Europe that is so prevalent along our roadsides. Linnaeus originally described the species in a different genus, and Koch later placed it in *Brassica*. *Erodium cicutarium*, a European weed in the geranium family, is introduced as well. Linnaeus also described this species as being in a different genus, and Aiton later explained that l'Héritier placed it in *Erodium*.

Correlative data may be derived by relating the locality of a particular collection with geographically indexed data (e.g., climate, soils, geology, hydrology). For example, if one knows the precise localities where *Rosa carolina* occurs, then one can infer site characteristics such as soil type, associated vegetation, annual precipitation, and so forth.

The original sources of information found in Floras are the observations and experiments made by botanists on the plants in natural habitats and their study of herbarium specimens. Botanists writing Floras also consult literature in which the results of previous studies have been summarized.

Information on plants, their characteristics and distributions, can be found in a vast number of published scientific research papers, which include revisionary or monographic studies on specific groups, and other floristic studies. The Flora of North America project in an incomplete survey has identified over 800 relevant monographic or revisionary articles published between 1970 and 1984, and hundreds of floristic projects of varying sizes and scopes are currently underway. Floristic information may also be found in ecological studies, which may encompass analyses of life-forms, phenology, relations to specific abiotic factors, form and function of plant structures, and descriptive analyses of plant communities.

Available sources of unpublished floristic information include Master's theses, Ph.D. dissertations, electronic databases, collectors' field notebooks, data files in local and regional herbaria, and notes compiled by private and public conservation agencies. Additional unpublished information, accessible only on an individual basis, remains associated with archival herbarium collections throughout North America.

---

[2] Herbarium specimens are pressed, dried, plant samples that are labelled with their names, general characteristics, and precise places of collection. Usually the plants with their labels are mounted on archival paper about 12" x 17", arranged systematically (alphabetically or taxonomically), and stored in cases. There are about 1,400 herbaria in the U.S. and Canada housing an estimated 60,000,000 specimens.

PART I: PRESENTATIONS

FLORA EUROPAEA AND THE EUROPEAN DOCUMENTATION SYSTEM[1]

-- **Vernon Heywood**

Professor Heywood was involved with the first FNA (Flora North America) program as an advisor, during a period when he was deeply enmeshed in the Flora Europaea project. At the time, he viewed the North American effort from the perspective of someone else concerned with producing a Flora. Now, however, his interests are mainly in conservation and the maintenance of genetic diversity, which makes him a critical consumer rather than a producer of taxonomy. Thus he is now able to assess floristic efforts and products, including his own, from both standpoints.

Flora Europaea was the first truly continental Flora. It dealt with 12,500 species, took 20 years to complete, and was published in five volumes between 1964 and 1980. From the start, there were significant problems to overcome. Plant taxonomy then was actually a very undisciplined field, with chaos just beneath the appearance of orderliness. There was great need for standardization, such as in terminology, but great reluctance on the part of individuals and institutions to embrace the compromises that would be required to achieve it. The contributors and editors had to cope with an old and very diverse European literature in 26 languages with 26 different traditions of taxonomic usage. Also, the European flora included many species whose typification was unresolved and difficult.

In the face of the many political and other sensitivities, certain basic policies were adopted at the outset and followed (enforced) rigorously. Treatments had to be consistent in their content and format. For this a Guide for Contributors (the "Green Book") was developed, and it has served as a model for others since. Keys had to work; that is, they would be designed for practical use rather than to convey taxonomic concepts. Finally, to help deal with the contents of the existing literature in an orderly and consistent way, the concept of "standard Floras" was developed. As a result of these policies and the necessarily collective nature of its operating procedures, the Flora Europaea project introduced the "collegium" into taxonomy, a precedent whose emulation has now become commonplace in floristic efforts such as Flora of North America. The essence of such collegial efforts is editorial oversight and review as a framework for individual work.

Editorial processing for Flora Europaea included five stages, involving review and editing of original manuscripts by regional editors, subsequent revision by authors, review of the revised manuscripts by regional editors, and finally review by the editorial committee before a final draft was prepared for the press.

---

[1]Reported and Compiled by John G. Packer and Robert W. Kiger

The approach to the Flora was a manual one--it was not a computer-based enterprise. Documentation of all sorts, such as of chromosome numbers and distribution data, was accumulated, maintained, and consulted on paper, in traditional files.

Cambridge University Press initially overestimated demand for the first volume when it printed 10,000 copies, but then underestimated demand for the second volume when it printed 5,000 copies. There had been no plan for further editions, but by the time the fifth volume had been published new printings were needed. Meanwhile the first volume had become seriously out-of-date, and royalties from sales, which were deposited in a trust fund, had accumulated to a substantial amount. So, the necessary resources being available, a decision was made to revise Volume 1, which was undertaken with a wholly different organization and procedure. When that is published in 1989, after 5 years of preparation, the project will finally be finished. This one revised volume will include a large number of changes, amounting to about 25% of the content, including 100 newly accepted taxa and complete overhauls of some groups, such as the ferns.

When the Flora Europaea effort was begun, concern was expressed in some quarters that the project would inhibit research in other areas of taxonomy. In fact, however, quite the opposite has occurred. The preparation and then the availability of the original five volumes actually stimulated much research of various sorts. A wealth of new information has resulted, and it became clear that a plan was needed to insure that the advance that Flora Europaea represents is not lost over the long term. The European Science Foundation sponsored a study that resulted in a Report on Taxonomy in Europe, which recommended setting up a permanent database for European floristic information.

To achieve such a database, an initial project called European Science Foundation--European Documentation System (ESFEDS) was funded in the amount of ?400,000 over five and a half years. The plan was to extract baseline data from Flora Europaea, add information on such things as chemical characteristics, design a database system to handle all these categories of information, and demonstrate its feasibility. There were many developmental problems, but the effort was ultimately successful, and the resulting system is able to provide valuable service for both research and reference. It runs using ORACLE software, with both mainframe and personal-computer capability, and it offers a user-friendly, ViewData-like interface on color monitors. The database and system can produce checklists for any area of Europe, complete with synonymies, and were used in preparing the revised treatment of ferns for the second edition of the first volume of the Flora. Also, they have been very useful in preparing a revision of European Orchidaceae. The initial ESFEDS project is now finished, and there has been some uncertainty about how and by whom it will be maintained as a permanent resource, thus to achieve its ultimate goal. No more project funding has been forthcoming, but recently an agreement was made with BIOSIS to take it over and assume responsibility for maintaining it and making it available.

It has become clear from the experience with Flora Europaea and ESFEDS that printed Floras, even when well-designed, are very difficult for non-taxonomists to use, and that databases with friendly interfaces are very desirable for providing general access to floristic information, including facilities for identification. Heywood is now working with the International Union for the Conservation of Nature and Natural Resources (IUCN), which has a conservation-oriented database containing records on

45,000 plant names, and uses those accepted in Flora Europaea as the standard names for European taxa (even though they are now not necessarily the correct ones, in light of more recent research). IUCN, together with the World Wildlife Fund and other international organizations, is now planning to establish a World Conservation Monitoring Centre. The results of projects like Flora Europaea and Flora of North America, especially when electronically accessible, will be valuable assets in the Centre's work, and for conservation efforts generally.

THE PRECIS COMPUTER SYSTEM AND ITS APPLICATIONS FOR THE FLORA OF SOUTHERN AFRICA

-- G. E. Gibbs Russell

INTRODUCTION

The Flora of Southern Africa project (FSA) is over 25 years old, and the PRECIS taxonomic computer system at the National Herbarium, Pretoria (PRE) was begun independently about a decade later. Since 1981, the two projects have been managed in parallel, so that information from the computer system aids authors in producing the conventional Flora, and so that PRECIS provides a continuously up-to-date overview of the entire southern African flora at shallow but gradually increasing levels of approximation, to provide plant information in the absence of a completed Flora.

The Flora of Southern Africa is a conventional "state of the art" Flora with keys, descriptions, synonyms, type citations, distribution maps, and line drawings. It covers the 24,000 species and infraspecies of plants from Botswana, Lesotho, Namibia, South Africa, and Swaziland. There are about 90 contributors, of whom about 40 live in southern Africa. Since the inception of the FSA in 1966 about 3,000 species (14% of the total) have been treated in published volumes. In addition, the FSA has led to coordination of taxonomic research in the subcontinent and has resulted in many supporting journal articles. Although successful in these minor objectives, the progress of the FSA has been unacceptably slow. However, this is an unavoidable result of the extremely high species diversity and the small taxonomic community in southern Africa. Computerization has proved a support and alternative for the conventional FSA treatment through providing the following: specific information to researchers for the FSA; distribution-based data not available in the FSA; a continuously up-to-date taxonomic treatment; and the basis for the identification service of PRE, which identifies 20,000 to 30,000 specimens annually.

PRECIS now consists of four separate but linked components for specimen, taxon, nomenclatural, and curatorial information. The "PRECIS experiment" covers an estimated 9% of the world's plant species. It provides a broad overview of the southern African flora as well as detailed information about particular species, specimens, and localities. PRECIS results have wide applicability in floristic, systematic, and vegetation studies, and could not have been obtained without computerization.

PRECIS development has followed the general trend of transferring data and functions from mainframe operation to microcomputers. The Specimen and (dormant) Nomenclatural components are presently restricted to a Burroughs 7900 mainframe whereas the Taxon component operates partly on the mainframe and partly on MS-DOS microcomputers. Curatorial-PRECIS works only on a Burroughs 28 microcomputer network. Data obtained from searches of all components can be provided as hardcopy, as ASCII or EBCDIC tapes, and on 5 1/4 inch diskettes for either BTOS or MS-DOS operating systems. Of the four

PRECIS components, Specimen and Taxon are most relevant to this workshop, and the others will not be discussed in detail.

PRECIS RESULTS AND THEIR APPLICATIONS

Specimen-PRECIS

The object of the specimen component of PRECIS is to make the data on herbarium specimen labels available in many different ways. This information provides a basic resource for our knowledge of the southern African flora, and meets specific needs of taxonomic researchers and herbarium curators. Data fields for each specimen include: collector code and specimen number, scientific name code, date, locality (region, quarter degree latitude and longitude grid reference, as well as major, minor, and precise locality), plant height, whether flowering or fruiting, altitude, aspect, habitat codes (biotic effects, vegetation type, substrate, moisture regime, soil type, life-form), herbarium code, and type or voucher code. The 640,000 specimens now included in Specimen-PRECIS represent about 30% of the ca. 2.2 million specimens thought to be held in southern African herbaria.

The Burroughs DMS database structure is now quite old, but it has proved sound, and the original special programming required for output not available under the DMS-INQUIRY search function has continued for eight years to produce the results we require. With recent additions to the special programs, Specimen-PRECIS can provide the following output in a standard format:

- Lists of all the specimens sorted by species, genus, and family, with the complete label information for each specimen. Within each species the specimens can be sorted by collector and number, or by locality, or by PRE accession number.

  Example application: Exchange of specimen data. Printouts of this information are filed in the Herbarium and copies are also provided to Flora-writers and monographers. In successful trials we have moved PRECIS specimen data to commercially available software (dBASE III+, DATAEASE, LOTUS 123) on microcomputers. Researchers correct the PRECIS information and augment it with specimens from other herbaria as their study progresses, and at the end will feed the complete and correct data back into Specimen-PRECIS.

- Checklists of all the species recorded for any area. The area may be defined as a single quarter-degree grid, as a region covering a number of quarter-degree grids, as an entire province or country, or as the whole of southern Africa. The total number of species and specimens recorded for the area are reported.

  Example applications:
  (1) Determining gamma diversity. Checklists for each biome compared by secondary programs gave the number of species and the number of endemics for each biome and the number of species shared between the biomes. This allowed us to analyze the relationships between the

biomes, revealing two major groups characterized by summer and winter rainfall respectively. Fynbos is the major winter rainfall biome, with Succulent Karoo closely related, and Savanna is the major summer rainfall biome, with Grassland, Nama-Karoo, and Desert closely related. Within these two major groups each biome is floristically distinct, each (except Desert) is rich in species, each is a center of diversity for certain large genera, and each (except Nama-Karoo) is distinguished by the occurrence of important families. These results should be used to guide conservation priorities.

(2) Distribution of higher taxonomic categories and correlation to other factors. Lists were generated of the grass subfamilies in each whole degree. This indicated the dominant subfamily in each region, which allowed correlations of the distribution pattern with simple environmental and physiological parameters such as rainfall season, rainfall quantity, and photosynthetic pathway. The resulting map showed that each subfamily dominates in the environment in which its photosynthetic pathway is most efficient: Arundinoids and Pooids (C_3 photosynthesis) in mesic winter rainfall areas; Chloridoids (C_4 aspartate photosynthesis) in arid summer rainfall areas; and Panicoids (C_4 malate photosynthesis) in mesic summer rainfall areas.

- Compilations of several checklists, with the number of specimens recorded for each species in the component checklists appearing in a matrix

  Example application: Production of a third edition of the *List of species of southern African plants*. Checklists of each of the eight major political territories (countries and provinces) were combined into a matrix. After checking in PRE and in other herbaria, publication of this list will provide definitive checklists for each territory as well as a coarse indication of the distribution of each species in southern Africa.

- Lists of all the quarter-degree latitude and longitude grids recorded for each species

  Example application: Species distribution maps. Maps can be either prepared by hand using the lists of grids, or computer plotted using the program DISSPLA operating under BTOS. An additional plotting program operating under MS-DOS is urgently needed.

- Lists of all the quarter-degree latitude and longitude grids in southern Africa, with the number of specimens recorded for each grid

  Example application: Specimen collecting intensity for southern Africa. Mapping the number of specimens for each quarter degree in size classes shows that the mesic east and south are far better known than the arid central and western regions. Since 1983, PRE has planned collecting trips using this information. The 1983 map showed 360

quarter degrees with more than 100 specimens, and the 1987 map shows over 900 grids with more than 100 specimens. The collecting intensity maps indicate the overall pattern of specimens that supports all floristic and systematic research on the southern African flora.

- Lists of the species in each "Wing" of PRE, with the number of specimens recorded for each species

  Example application: Monitoring and controlling accessions to PRE. Special programming in Curatorial-PRECIS allows checking the number of specimens recorded for each species in each quarter-degree grid. All specimens from undercollected areas are retained, but for species with more than 10 specimens already held from a particular quarter degree, no new accessions are taken up from that quarter degree unless they are types, vouchers, or are otherwise outstanding for completeness or good label data. PRECIS thus has as a feedback mechanism to control its own growth through guiding the growth of PRE. This limitation in specimen numbers is necessary because of storage space restrictions in the Herbarium.

- Lists of all the collectors who have contributed specimens to PRE, with the number of specimens attributed to each

  Example application: Provide the basic list for a biographical encyclopedia of plant collectors. Gunn & Codd (1981), *Botanical exploration of southern Africa*, treated all collectors with more than 200 specimens recorded in PRECIS.

- Collector's registers with complete label information for specimens in PRE

  Example applications:
  (1) Personal collecting registers. Any collector is provided upon request with a register of all his collections in PRE. This not only forms a record of his localities, but can inform him of the current identification of his specimens.

  (2) Tracing routes of important collectors. Many important early collectors did not leave adequate records of the localities they visited. The PRECIS registers can be sorted by date or by specimen number to show the route traveled on collecting trips.

- Flexibly designed searches, through DMS-INQUIRY, sort any of the 24 data fields and report them, specimen-by-specimen, in any order

  Example applications:
  (1) Planning field study of a genus. Lists of all specimens in each species sorted by quarter-degree grid and separately sorted by date of

flowering or fruiting will indicate the optimal season and route to collect all species in the required phenological state.

(2) Determining phenology of economically important plants. Retrieving flowering times for important pasture grasses in Natal showed that there were early-flowering, early-and-late-flowering, and late-flowering species. When correlated to seral state and pasture value, it was found that early-flowering grasses are of low seral state and are unpalatable, early-and- late-flowering grasses are of intermediate seral state and are highly palatable, and late-flowering grasses are generally of the highest seral state and are unpalatable.

(3) Finding species to be included in specific studies. A list of all climbing plants recorded from the Drakensberg mountains of the Orange Free State, Natal, and Lesotho was prepared as a beginning for a field guide to these species.

Future growth of Specimen-PRECIS will continue to center on adding new accessions to PRE and improving the quality of data on each specimen. Further developments will include links to other herbaria and development of an MS-DOS database to handle Specimen-PRECIS data for researchers.

The major consequence of adopting a herbarium specimen system is the absolute necessity for accurate names on all specimens entered. When this ideal is not achieved, the usefulness and accuracy of a specimen system are diminished. The commitment to accurate identification is expensive because a continuous curatorial input in both herbarium and computer is required to maintain it. When two or more herbaria share a common database, firm management agreements are necessary in order to maintain adequate quality of plant identifications.

Taxon-PRECIS

The object of the Taxon component of PRECIS is to provide a continuously up-to-date inventory of the families, genera, species, and infraspecific taxa of plants in southern Africa at gradually increasing levels of approximation. Taxon-PRECIS thus gives an overview of the flora that can guide research, and that can be easily consulted for information about plants, particularly by non-taxonomists. Taxon-PRECIS is complementary to the FSA, whose completion is inevitably a long-term goal. In the meantime, Taxon-PRECIS provides complete coverage of the entire flora at a shallower level, with continuous incorporation of the most recent treatment of individual plant groups.

The taxon database originally planned for PRECIS was a simple one relating species names through codes to genera, families, common names, and specimens. However, a number of the 54 data fields in the original specimen database recorded descriptive characteristics and economic properties, information that is relevant to taxa rather than to specimens. This confusion between specimen and taxon data was partly responsible for the complexity of the first specimen database, which led to its downfall.

When the specimen database was converted to its present structure in 1981, all "taxon" data were eliminated. The early inclusion of taxon data in PRECIS

reflected a real need for information about plant species that is otherwise unobtainable in a region lacking a completed modern Flora. This need has been met in two ways. First, much of the special programming in Specimen-PRECIS is designed to accumulate and summarize separate specimen records so that they resemble taxon data, for example, the various checklist options, and the lists of species in each herbarium "Wing" mentioned above. Second, the need for information about plant species led to the separation of the present Taxon-PRECIS from Specimen-PRECIS during the first year of the new system. Since then, two levels of approximation have been completed in Taxon-PRECIS, and information is being added for the third approximation.

- Approximation 1. In 1982, the current "Wings" listings of names applied to specimens were obtained from Specimen-PRECIS as data files. These were exhaustively checked by the staff of PRE to determine the currently accepted scientific name for each taxon, to determine which taxa were naturalized components of the flora, and to ensure correct spelling and consistent author abbreviations. Names from literature were added for taxa not represented by specimens in PRE. Trial listings were circulated to other southern African herbaria who also made changes and additions. The resulting data files provided the first treatment of the southern African flora at species level.

  Example application: Basic regional checklist. The completed data files were published as the *List of species of southern African plants*. In addition, the corrected data were fed back to Specimen-PRECIS to correct the names on specimens. Analysis of "Approximation 1" led to the first accurate count of the taxa occurring naturally in southern Africa, and allowed internal and external comparisons of the largest families and genera. It was shown for the first time that, with 24,000 species and infraspecific taxa in 1,700 genera and 320 families in an area of only about 2,500,000 km^2, southern Africa probably has the richest temperate flora in the world.

- Approximation 2. While determining the currently accepted name for each taxon in "Approximation 1", it became obvious that the information in a bare list of names was insufficient because of the large number of name changes, especially in recent years. Therefore, recent synonyms and the literature references needed to establish the synonymy and to identify species in each genus were added to the "Approximation 1" data files. After the list was completed, additions, deletions, and alterations based on current literature were added annually.

  Example application: Augmented and continuing regional checklist. The data files were published as the second edition of the *List of species of southern African plants*, and the additions and alterations have been published annually in the journal *Bothalia*. Analysis of the data files has shown for the first time the number and categories of name changes recorded for our flora in recent years. During the four years for which

annual records have been kept, over 7% of the taxa have experienced name changes. Of these changes, only 9% have been nomenclatural changes necessary under the ICBN, while 91% of the changes have resulted from taxonomic reclassification at generic or specific level. Given the extremely large number of plant groups not recently studied in detail; it is obvious that there will be many name changes in the future as more groups are revised.

- Approximation 3. The aim of this approximation is to record commonly needed information for all species. Possible fields have been determined from a number of sources: kinds of information included in the *Flora of southern Africa*; data available in Specimen-PRECIS; comparison with other taxon-based computer systems; and consultation with potential users in southern Africa. Most of the possible data fields correspond to the "Type 1" data for ILDIS (International Legume Database and Information Service): accepted name and author, synonyms with references, vernacular names, suprageneric classification, habitat, distribution, life form, conservation status, importance to man, references to descriptions, illustrations, and maps, and notes. Local taxonomists recommend that brief diagnostic descriptions, flowering times, and voucher specimens also be included.

Data to be transferred from Specimen-PRECIS include distribution, habitat, and flowering time. In preparation for transfer to Taxon-PRECIS, work began in 1985 on an extensive program to improve the quality of locality data represented by quarter- degree grid references in Specimen-PRECIS. Because this will be a lengthy project, it was begun long before the data fields or data structure for "Approximation 3" were considered. A useful by-product will be a revised gazetteer for southern Africa that includes quarter-degree grid references for all plant collecting localities in Specimen-PRECIS.

So far, Taxon-PRECIS has been deliberately maintained as simple data files manipulated by simple programs that aid data entry and retrieval. There are two reasons for this: firstly, dissatisfaction with the rigid structure, limited output options, and machine-bound nature of the Specimen-PRECIS database made us reluctant to establish another database; secondly, since developments in taxon-based systems were taking place in other centers, it seemed wise to maintain the computer treatment in a relatively primitive state so that we can establish compatibility with similar systems as they develop elsewhere.

As an experiment in appending other fields to the present data files, illustration information was added for the ca. 2,000 species treated in the first 49 volumes of *Flowering Plants of Africa*. Difficulty with input to the very large data files indicates that additional information will require a more complex data structure. The basic decision will be whether to continue Taxon-PRECIS as program-manipulated data files or to redevelop it as a database. Fortunately, software developed elsewhere is available for trial. A prototype using ALICE, a database application of dBASE III, is planned for legumes to work in parallel with ILDIS. A prototype using DELTA programs is already established for grasses and works in parallel with Leslie Watson's project on world grass genera [Research School of Biological Sciences, Australian National University, Canberra]. The final decisions about data fields

and data structures for the whole of Taxon-PRECIS will be made after comparing and evaluating the completed prototypes. In the meantime, we will continue to operate "Approximation 2" and to accumulate accurate distribution records.

Example applications:
(1) Index to *Flowering plants of Africa* plates. All names were extracted to form an index to the 50-volume series, beginning with the first volume in 1922. The index gives the currently accepted name and relevant synonyms for each species included. For the first time, PRECIS data for publication was submitted directly to the printer electronically, on 5 1/4 inch diskette.

(2) Identification of grasses. The *Manual of the grasses of southern Africa*, treating 205 genera and 970 species and infraspecific taxa, is being prepared directly from data files under DELTA for the generic keys and the generic and species descriptions. Distribution maps are prepared directly from Specimen-PRECIS. Diskettes operating under MS-DOS will be distributed with the published volume, making all the descriptive and distribution data available for electronic data retrieval and specimen identification using the DELTA program INTKEY.

Maintaining even a simple taxon-based system requires a commitment to keeping it up-to-date for new taxa, for new classifications, and for additional data in all the fields. This can be done only through constant scanning of literature and herbarium accessions. Finally, the potential alterations and additions discovered must be evaluated by experts before being entered in the system.

Nomenclatural-PRECIS

The object of the nomenclatural component of PRECIS is to speed up the progress of taxonomic study for the *Flora of Southern Africa* by providing the bibliographic information needed to evaluate each scientific name based on or applied to southern African plants, in a form that can be sorted easily for locating literature and type specimens.

Unlike the other three PRECIS components, Nomenclatural-PRECIS was not part of the original system. However, it is of similar age, the first versions having been planned as a cardfile "Master Index" to plant names similar to that outlined by N. T. Burbidge for the Australian flora. Considerable work was done in the Poaceae, and in 1979 computerization of the existing cards as well as capturing additional names began for this family, using program-driven data files. A manuscript list of all names for Mesembryanthemaceae was completed in the same year. These two families, which have similar numbers of species and names but differ greatly in distribution, ecological and economic importance, and taxonomic history were compared to assess possible future growth of the system. Both families have about four names per taxon. Since there are 24,000 taxa to be covered for FSA, the estimated minimum size for a completed Nomenclatural-PRECIS is about 96,000 names.

A. D. Chapman (pers. comm.) reports that, after 14 years' work, the completed Australian Plant Name Index contains about 60,000 names. Burbidge had

estimated a three-year term for the project. Completion of Nomenclatural-PRECIS would take longer because more names must be included. Work has been suspended because we lack the resources to continue the project at this time. Should work on Nomenclatural-PRECIS be resumed it would appear most efficient to adopt the successful Australian software and obtain names not already included in Taxon-PRECIS from other computerized indices, notably *Index Kewensis*.

> *Example Application:* Grass names. Data for Poaceae up to 1980 includes genus, species, and infraspecies names; authors; publications; dates; basionyms; other genera to which an epithet may have been referred; and types (collector, specimen number, locality, and herbarium location if known). These data may be sorted either by genus and species or by publication. The data remain available for search, and a few requests are answered each year.

Curatorial-PRECIS

The objects of the curatorial component of PRECIS are to perform routine herbarium tasks more accurately and efficiently through computerization of repetitive clerical work, to make use of information in the Specimen and Taxon components to guide the growth of PRE, and to serve as the interface of PRECIS in the herbarium by acting as the primary data entry point for specimen data and some kinds of taxon data.

Curatorial-PRECIS is the youngest offspring of Specimen-PRECIS. It began as a separate system in 1986 and is only partly complete. Four factors led to its origin: (1) dissatisfaction with the slow delivery of specimen labels from Specimen-PRECIS; (2) the realization that the basic list of species names, transferred from Specimen-PRECIS, could be accessed through code numbers to enter correctly spelled names on specimen labels, *determinavit* lists, and loan lists; (3) the need for herbarium staff to have direct access to computer terminals in order to use and enter information in the Specimen and Taxon components; (4) the need to develop an in-house system on microcomputers to handle information brought down from Specimen- and Taxon-PRECIS and thus avoid paying heavy computer-time charges on the mainframe.

This component of PRECIS is still under active development, and it is anticipated that considerable changes will take place in linking the older components once it is fully operational. Parts of Curatorial-PRECIS functioning so far are a prototype system for specimen labels, a central address file that links most parts of the curatorial system, control of the identification service, compilation of *determinavit* lists, preparation of loan lists for outgoing specimens, and record-keeping for name changes that interfaces with both Specimen- and Taxon-PRECIS.

CONCLUSIONS

PRECIS has given us access to detailed specimen, species, and locality information that has provided a unique overview of the southern African flora. The consequences of this overview profoundly affect taxonomic research in southern Africa, and also reach beyond taxonomy to more applied disciplines.

1. We know the strengths and weaknesses of the herbarium specimen coverage that underlies all taxonomic research. We know where to direct field study, and we can calculate storage space and curatorial needs for accessions resulting from a planned collecting program.
2. We have a complete and continuously up-to-date coverage for the southern African flora at species level. This information is accessible to all who use plant names or require information about plants.
3. We know the size of our flora. The actual count is about 25% higher than previous estimates. This has obvious implications for the funding and duration of the Flora of southern Africa project. Furthermore, the extremely high species diversity must be considered as an ecosystem attribute important for management and conservation of vegetation in the subcontinent.
4. We know the floristic relationships between the six major biomes and we can generate checklists for any area. This information must be used in setting conservation priorities.
5. We can measure, in costs and man-hours, the effort required to keep our species treatment up to date according to the International Code of Botanical Nomenclature.
6. We have an abundance of data to test biological hypotheses, for example through character mapping, phenological studies, and distribution of higher taxa correlated with environmental factors.
7. We can control the number of specimens accessioned by PRE, accepting for the most part only specimens from areas poorly represented in the herbarium.
8. We have an efficient plant identification service that can produce *determinavit* lists on demand. It is no longer necessary to proofread the scientific names and authorities.
9. We can effectively monitor and manage herbarium curatorial functions.

The unique PRECIS experiment in computerized coverage of a major flora and computerized procedures in a large and active herbarium accounts for about 9% of the world's plant species. The results so far have well repaid the effort and the diversion of labor from other worthwhile projects. The PRECIS story has included errors, faulty planning, redoing, patching, and carrying on regardless of problems. The success of the system in the long term has resulted from continual dissatisfaction with it at any one time: "What we have now works OK, but it would be marvelous if we could only...."

ACKNOWLEDGMENTS

The PRECIS team has never been large. Dr. B. de Winter, Director of the Botanical Research Institute, has promoted PRECIS from its very beginning. Over the years, data officers in charge of the system have been J. W. Morris, R. E. Magill, and G. E. Gibbs Russell; programs have been written by J. W. Morris, B. Young, W. Kruger, and B. C. de Wet; database managers have been H. F. Glen, P. Gonsalves, N. P. Barker, and J. C. Rautenbach; technical assistants have been J. C. Jooste (deceased), E. Evenwel, H. von Ronge, and W. Roux; and data input has been done by H. P. van der Westhuizen and J. Mulvenna.

PERTINENT LITERATURE

Edwards, D., and O. A. Leistner. 1971. A degree reference system for citing biological records in southern Africa. Mitt. Bot. Staatssamml. Muench. 10: 501--509.

Gibbs Russell, G. E. 1983a. Correlation between evolutionary history, flowering phenology, growth form, and seral status for important veld grasses. South African Journal of Botany 2: 175--180.

--------. 1983b. Register of names and types in Poaceae: a computerized index for southern Africa. Bothalia 14: 943--944.

--------. 1985a. PRECIS: the Botanical Research Institute's computerized information system. S. Afr. J. Sci. 81: 62--65.

--------. 1985b. Analysis of the size and composition of the southern African flora. Bothalia 15: 613--630.

--------. 1986. Significance of different centers of diversity in subfamilies of Poaceae in southern Africa. Paleoecology of Africa 17: 183--191.

--------. 1987. Preliminary floristic analysis of the major biomes in southern Africa. Bothalia 17: 213--228.

--------. 1988. Distribution of subfamilies and tribes of Poaceae in southern Africa. Proceedings of the Eleventh Plenary Meeting of AETFAT. Monographs in Systematic Botany from the Missouri Botanical Garden No. 25.

--------, and H. F. Glen. 1984. Register of names and types: a comparison between Mesembryanthemaceae and Poaceae. Bothalia 15: 125--130.

--------, and P. Gonsalves. 1984. PRECIS - a curatorial and biogeographic system. *In* R. Allkin and F.A. Bisby, *Databases in Systematics*, The Systematics Association Special Volume No. 26. Academic Press, London.

--------, G. Germishuizen, P. Herman, P. Olivier, S. M. Perold, C. Reid, E. Retief, L. Smook, J. van Rooy, W. G. Welman, and P. Gonsalves. 1984. List of species of southern African plants. Memoirs of the botanical survey of South Africa 48.

--------, C. Reid, J. van Rooy, and L. Smook. 1985. List of species of southern African plants, Ed. 2, Part 1, Cryptogams, Gymnosperms, Monocotyledons. Memoirs of the Botanical Survey of South Africa 51.

--------, C. Reid, L. Fish, G. Germishuizen, M. van Wyk, J. van Rooy, and Staff. 1987a. New taxa, new records, and name changes for southern African plants. Bothalia 17: 269--275.

--------, E. Retief, and L. Smook. 1984b. Intensity of plant collecting in southern Africa. Bothalia 15: 131--138.

--------, W. G. Welman, E. Retief, K. L. Immelman, G. Germishuizen, B. J. Pienaar, C. M. van Wyk, and A. Nicholas. 1987. List of species of southern African plants, Ed. 2, Part 2, Dicotyledons. Memoirs of the Botanical Survey of South Africa 56.

--------, W. G. Welman, G. Germishuizen, E. Retief, B. J. Pienaar, C. Reid, L. Fish, J. van Rooy, C. M. van Wyk, E. Kalake, and Staff. In press. New taxa, new records, and name changes for southern African plants. Bothalia.

Gunn, M., and L. E. Codd. 1981. *Botanical exploration of southern Africa*. Balkema, Cape Town.

Killick, D. J. B., and E. du Plessis. 1988. Index to volumes 1 - 49. *Flowering Plants of Africa*.

Magill, R. E., G. E. Gibbs Russell, J. W. Morris, and P. Gonsalves. 1983. PRECIS, the Botanical Research Institute herbarium data bank. Bothalia 14: 481--495.

Morris, J. W. 1974. Progress in the computerisation of herbarium procedures. Bothalia 11: 349--354.

--------. 1980. Encoding the National Herbarium (PRE) for computerized information retrieval. Bothalia 13: 149--160.

--------, and H. F. Glen. 1978. PRECIS, the National Herbarium of South Africa (PRE) computerized information system. Taxon 27: 449--462.

--------, and O. A. Leistner. 1976. Southern african place names. Ann. Cape Prov. Mus. 12: 1--565.

--------, and R. Manders. 1981. Information available within the PRECIS data bank of the National Herbarium, Pretoria, with examples of the uses to which it may be put. Bothalia 13: 473--485.

Pankhurst, R. J. 1984. A review of herbarium catalogues. *In* R. Allkin and F. A. Bisby, *Databases in Systematics*, The Systematics Association Special Volume No. 26. Academic Press, London.

Staff of the National Herbarium. 1985. New taxa, new records and name changes for southern African plants. Bothalia 15: 751--759.

Staff of the National Herbarium. 1986. New taxa, new records and name changes for southern African plants. Bothalia 16: 109--118.

FLORA OF AUSTRALIA

-- **Alex George**

INTRODUCTION

The Australian Biological Resources Study

Recognizing that lack of funding had resulted in a poor international record by Australia in biodiversity (flora and fauna) studies, the Australian Biological Resources Study (ABRS) was initiated in 1973 under an Interim Council. This Council and its successors stimulated taxonomic and ecological studies of Australian biodiversity through a grants scheme and has begun publication of major national series on flora and fauna.

ABRS provides advice to the Australian Government on national taxonomic collections, establishes and maintains taxonomic databases, coordinates vegetation mapping, and publishes flora and fauna handbooks. Much of the research and documentation work of the ABRS is by staff of museums, herbaria, and universities. Some Divisions of the Commonwealth Scientific and Industrial Research Organization are also involved in ABRS-related research. ABRS is an excellent example of cooperation between the Commonwealth, States, and the scientific community under the leadership of the Minister.

ABRS aims to answer two questions: What kinds of plants and animals are there in Australia? Where are these plants and animals found? In detail, its objectives are to:

- Coordinate all work aimed at collecting, describing, and classifying Australian plants and animals and determining their distributions at a national level
- Improve and increase the national taxonomic and biogeographic effort
- Develop the Bureau of Flora and Fauna as a national and international resource center for information on biogeographic and taxonomic research and documentation in Australia
- Establish priorities for funding taxonomic research by the Australian Government to enable regular publication of a systematic series of flora and fauna handbooks
- Maintain information on the scope and status of Commonwealth taxonomic collections

Geographically the activities of ABRS encompass all States and mainland Territories, immediate offshore islands, Macquarie Island, Lord Howe Island, and all external Territories.

Background to the Flora of Australia project

Australia has had only one national flora--*Flora Australiensis*, written by George Bentham and published in seven volumes between 1863 and 1878. Bentham worked exclusively in Europe, mainly at Kew, and his *Flora* was based entirely on herbarium material. He covered 8,125 species of vascular plants, including the ferns.

At that time botanical exploration of the continent was little more than a century old. Large areas were unexplored, the population was small, and indigenous botany was in its infancy. The number of taxonomic botanists has always been small, and even today there are only about 100 (some retired but active) in Australian herbaria, universities, and other organizations. Today, too, there are still large areas that are botanically unexplored or little-known.

The number of known taxa in our flora is thus increasing with both exploration and refinements in the taxonomy, and it is likely to be another century before a reasonably complete picture emerges. Nonetheless the need for a new Flora has been increasingly felt as *Flora Australiensis* became more outdated. The first public call came as long ago as 1910. It was not until the late 1950s, however, that a concerted move began to convince the Commonwealth Government of the need for a new Flora.

Support for the project came from a number of sources. Taxonomic botanists were acutely aware of the need, but beyond discussing it among themselves there was no one prepared to take it further until a newcomer to the scene began sowing the seed. Hansjoerg Eichler, appointed keeper of the State Herbarium of South Australia in 1954, actively promoted the concept among colleagues, and in 1958 it was discussed in depth at the 33rd Australia-New Zealand Association for the Advancement of Science (ANZAAS) Congress. ANZAAS set up a Flora of Australia Committee, and through this and other avenues continually pushed for the project.

The matter was taken up by the Australian Academy of Science, which broadened the concept to a biological survey of Australia. The Academy also gave the concept continued support until it was finally adopted by the Government.

Among a number of amateur societies involved with the Australian flora, the Society for Growing Australian Plants was a vigorous campaigner. Faced with trying to identify plants when there was no national Flora and few regional ones, the Society's members were acutely aware of the potential value of a new work. Since publication began in 1981, the Society's members have provided a significant market for the *Flora*. Because of small numbers and large distances, Australian taxonomists had not communicated well, but in 1973 they formed a professional group, the Australian Systematic Botany Society.

The Society served to bring people together, fostered more communication, and prepared an index of current plant taxonomic research. Under its auspices a *Flora of Central Australia* was written and published. This both directed attention to the need for modern Floras and gave some 50 people experience in Flora-writing. The Society now has over 300 members. In Australia during the 1960s there was great expansion in agriculture, mining, and industry. Interest in and concern for the environment grew in parallel, with the corresponding increased demands for information on the flora and fauna. By the early 1970s environmental concern was

sufficiently widespread for both major political parties to include proposals for a biological survey in their platforms for the 1972 Federal Election.

The incoming Labor Government acted quickly to set up the Australian Biological Resources Study in 1973, initially with an interim committee whose brief was to advise the Government on how the Study should be developed. The committee's report was referred to various other committees, leading in 1978 to the formal establishment of ABRS within the Department of Science. A new *Flora of Australia* was one of its objectives. An ABRS Advisory Committee was appointed, and the formal announcement of the Flora project was made on 24 April 1979.

The final administrative step in setting up ABRS was taken in November 1979 when the Bureau of Flora and Fauna was established to provide staff to put the program into effect. Over the next few years, staff were appointed, the Flora Section eventually having five staff, three of whom work almost full time on the *Flora of Australia*.

In October 1979 a further advisory body was appointed, the Flora Study Group, later to become the Flora Editorial Committee. This body advised the Government on the scope, systematic arrangement, and format for the *Flora*, and is still retained for this purpose as required. The Committee does not have an active role in the day-to-day work of organizing, editing, and publishing the *Flora*. It consists of six voting members plus a Chairman, who has a deliberating vote. They are selected for their experience and expertise in plant taxonomy and are normally appointed for a three-year term. A broad geographical spread across the States is also maintained. The Director of the Bureau, the Assistant Director (Flora), and the Executive Editor are ex-officio members of the Editorial Committee.

DESCRIPTION OF THE *FLORA OF AUSTRALIA* PROJECT

Coverage

The *Flora of Australia* will include both the vascular and the non-vascular plants. We estimate around 20,000 vascular plants. These will be covered in the first 48 volumes (averaging 300 pages each). Volume 49 will cover two islands in the Tasman Sea--Lord Howe and Norfolk Islands--and Volume 50 the other oceanic islands. Volumes 51--53 will be the bryophytes (ca. 1,500 species), and Volumes 54--58 the lichens (ca. 2,000 species). We have not yet planned the algae and fungi, which could take another 20--30 volumes.

The Role of the Flora Editorial Committee

Among matters discussed by the Editorial Committee at early meetings were the following:

- Taxonomic arrangement. A modern phylogenetic system was adopted and the *Flora* divided into volumes. All families should be covered, even if only as status quo treatments. Arthur Cronquist's system, then almost ready for publication in its latest revision, was selected. Initial planning was restricted to the vascular flora.

- Readership. The *Flora* should be useable by a wide range of people having at least a fundamental biological knowledge. Use of highly technical terms would be kept to a minimum.

- Format. Published Floras and sample treatments already prepared formed the basis of discussion, which covered all aspects including descriptions, bibliographic data, distributional data, specimen citation, maps, key format, and illustrations. All details were addressed, down to punctuation and typeface.

- Authors. The Committee recommended authors for Volume 1 and early taxonomic volumes, and maintains this role.

A major advantage of this planning has been that the editorial staff has basically had the task of putting the program into practice. Inevitably other editing matters arise, but mostly these are minor. It has also the advantage that the format is not seen as that of an individual, a situation that could involve the Editor and staff in numerous arguments with contributors. The Editorial Committee recommended that, partly for this reason and partly for the sake of consistency throughout the *Flora*, the Editor should not leave a strong personal stamp on the work. One of the Committee's terms of reference is to act as arbiter in disagreements between the Editor and contributors. In the first seven years they have been asked to consider only one such matter.

Commencement of the *Flora* came at an appropriate time for decisions on standard abbreviations for authors, journal titles, and book titles. The Draft Index of Author Abbreviations compiled at the Royal Botanic Gardens, Kew, was published in 1980; Botanico-Periodicum-Huntianum had been published in 1968, and Stafleu and Cowan's Taxonomic Literature 2nd edition had begun to appear in 1976. All provided systems that, while not in all cases satisfactory, were available and comprehensive. For the sake of simplicity it was decided to adopt these with no exceptions.

On the matter of style I should point out that as a Government department we are required to publish through the Australian Government Publishing Service (AGPS), who provide advice on all publishing matters. The service has its own highly respected Style Manual to which Government departments are expected to adhere. In the case of the *Flora of Australia* there were several instances where botanical practice differed from AGPS style, e.g., literature abbreviations. These matters were fully discussed with AGPS design staff who agreed to accept them in the *Flora*. Indeed, AGPS were closely involved at all stages of planning format. We typeset the Flora in-house, and since the first two or three volumes were prepared there has been little need to involve AGPS design further, while they in turn feel no need to check our typeset text in detail.

In terms of content, the *Flora of Australia* is a compromise, as indeed any work must be that is prepared by more than one person. I believe that the strong support so far given the *Flora* is due in part to the fact that

- most people accept that it must be a compromise

- those involved in its planning have been highly respected

- people are prepared to accept the majority decision even if it is not their preferred option

The Editorial Committee members in particular have taken this approach. Each argues a particular viewpoint but, when a decision is made, accepts it and is ready for the next matter. The Committee has also adopted the principle that all matters, once decided, should not be re-opened unless there is a strong reason.

Content

Probably the most important compromise affecting the Flora has been on the level of detail. The possible extremes were a monographic work and a very concise, status quo work. What we have lies between. Thus, we include full bibliographic data on accepted names, basionyms, and synonyms, including type collections and where they are housed. All names that have been used for Australian plants--native and naturalized--are included, even if they must be listed as doubtful. If a name has been misapplied then it is cited with an example of the misapplication.

The *Australian Plant Name Index* (APNI) provides the bibliographic backup for these data. This Index has been compiled from primary sources and includes all names from the level of genus down for the vascular flora. It is thus a marked advance over *Index Kewensis*. Currently the APNI contains ca. 60,000 names. It is the only part of the Flora program set up as a database. We also cite several published illustrations for each species. Preference is given to those in recent or readily available works.

The length of descriptions is limited to 150 words for families and 100 words for genera and species. This policy has worked reasonably well for containing the size of the Flora. It is a recurring matter for discussion, however, with authors who consider their groups of special importance or complexity and in need of longer descriptions. We have become adept at reducing overlength descriptions and so far, through this and discussion, have reached agreement with contributors on meeting the limits.

Distribution is described in standard geographical terms, avoiding local names that, although meaningful to people in the region, are not found in official maps. In the same paragraph habitat data are given. Distribution is supplemented by listing occurrence in a standard set of semi-natural regions recently published. A small map is also provided for each species or infraspecific taxon. Although the scale is small and the maps therefore somewhat crude, they have proved useful to many users.

A somewhat controversial issue in an essentially concise Flora is the inclusion of specimen citation. It is limited to either a maximum of five collections or, for taxa found in all Australian states, one from each state and territory. The criteria are left to the contributor but include morphological and geographical range. A discussion usually completes the species treatment. This is usually in telegraphic style and may cover such aspects as diagnostic characters, variation, relationships, usage, and other points of interest. Many early Flora contributors did not include this text, apparently in the mistaken belief that we required only the standard

treatment. It has taken some communication to spread the word that we encourage what is one of the most interesting and useful parts of the text.

Keys are indented and dichotomous. Once the decision on key format was made there has been little subsequent discussion or difficulty with this part of the text.

We allow descriptions of new taxa, lectotypifications, etc., in an Appendix to each volume. The one stipulation is that the treatment must be concise. Where a contributor has only a few entries and is not publishing a major paper beforehand, we encourage use of this section. It avoids any uncertainty over whether or not the names will be published in time for the relevant Flora volume.

A detailed Guide for Contributors is provided to each contributor.

Illustrations

Both line illustrations and color plates (photographs) are included in the Flora. For line illustrations the aim is to illustrate each genus, and, within genera of several to many species, a selection of species. We average one page of figures to 4-6 pages of text. Preparation of figures is arranged in consultation with the contributor. Several Australian herbaria have an illustrator on the staff who may be available to prepare Flora figures. A very few taxonomists draw their own. In other cases an artist must be contracted and is paid from Bureau funds. Currently we pay A$150 per page and have A$4,000--5,000 per year for such work.

To assist artists we have prepared a Guide for Illustrators. This explains our basic requirements and includes examples of good and poor figures. Both this Guide and the companion Guide for Contributors are available free.

A small selection of color plates is included, using color slides. These are placed four to a page in the introductory section. They were included mainly at the request of AGPS, to improve the popular appeal. Usually a color painting features a representative species from a volume on the cover and frontispiece. These are prepared by invitation for the Flora, and are later bought by Artbank, a Government agency that hires out works of art to other Government departments.

Timescale

When the project began, a publication rate of three volumes per year was envisaged. This would have seen the vascular Flora completed in 15--20 years. We now see that this was a gross overestimation, since we have published nine volumes in 7.5 years and with the most efficient use of present resources could not manage more than two per year. The reasons are several, and will be discussed further below, but the most significant are manuscripts in poor format requiring major basic editing/rewriting; failure to meet deadlines; insufficient editorial staff; underestimating the time needed for editorial planning, especially the non-vascular groups; and lack of funds to support Flora writers, travel to herbaria, and field work.

PRODUCING THE FLORA OF AUSTRALIA

Editorial Staff

Most of the preparatory work for the Flora of Australia on format, systematic arrangement, possible contributors, etc., was done by the Flora Editorial Committee. It is the task of the Bureau of Flora and Fauna to produce the Flora. Within the Bureau, the Flora Section has a staff of five: an Assistant Director, an Executive Editor, a Flora Writer, a Bibliographer, and an Editorial Assistant.

The Assistant Director's duties are mainly administrative, while the Bibliographer's main duty currently is compiling the Australian Plant Name Index. The Index, even though not yet complete, is already an essential tool for every plant taxonomist in the country. The Bibliographer checks all bibliographic data in Flora manuscripts against APNI and currently assists with keyboard work when the text is being typeset. The Editor's duties cover detailed planning, coordination, final manuscript checking, and supervision of typesetting and publication. This includes much of the correspondence with contributors and referees, though some is delegated. With a small staff it is quite a juggling act to assign tasks in such a way as to keep the program moving. Although we try to organize tasks in such a way that, when one of the staff is away on leave or travel, another can step in, this can only mean a slowdown elsewhere. The Editor may contribute to the Flora text, but time for this is very limited.

The Flora Writer's duties are both interesting and at times frustrating. They are essentially to write those taxa for the Flora for which no contributor can be found. This means that a wide range of taxa is studied, but that intriguing problems must often be left for other workers. Because of the need to complete manuscripts on a firm schedule there is simply no time for such diversions.

Our Flora Writer has one major management role--that of arranging and overseeing the preparation of illustrations (line drawings). Some institutions have a botanical artist on staff who can prepare figures for contributors at those institutions. A few contributors have both the expertise and the time to do their own illustrations. In other cases the Bureau arranges for a freelance artist to work in conjunction with a contributor, though the funds available for this work are limited. We maintain a register of artists interested in this work.

The main duty of the Editorial Assistant is to check new manuscripts for format, consistency, completeness, and accuracy. A new manuscript is entered on disk at the Bureau and from then until completion of typesetting is edited on a word-processor. New manuscripts are accepted either on compatible disks or as printout suitable for an optical character reader. The Assistant also makes up the pages of figures and maps and shares the keyboard work when the text is being typeset.

With other duties that inevitably arise in a Government department, we effectively have the equivalent of three staff working full-time on the Flora.

A word here on the Australian Plant Name Index (APNI). The project was started in 1973 under the Australian Biological Resources Study. It lists all names from the rank of genus downwards that have been used for Australian plants both native and introduced. The names have been checked against primary sources, not taken from lists such as Index Kewensis. In listing infraspecific and infrageneric names it is also more comprehensive than other indexes. The data include the place and date of publication, type citation, location of the type, family, listing of

relevant revisions since 1935, and comments if necessary. The APNI will be issued in microfiche form in about two years' time. It has been widely publicized, however, and most taxonomists now consult it when commencing a new group. All Flora manuscripts are checked against the APNI for accuracy of bibliography and type data.

Recruitment of Contributors

As mentioned before, a function of the Editorial Committee is to recommend possible contributors. There are four main sources: (1) taxonomists in herbaria, universities, etc., who can incorporate the work into their research program; (2) retired taxonomists; (3) freelance botanists funded by ABRS grants; (4) the in-house Flora-writer.

Usually it is our practice to invite an expert in the group to prepare the Flora text. Occasionally there is a choice, but with Australia's small botanical community this is rare. While we variously try to persuade someone to take on the task, we are never too insistent, for an unwilling contributor is likely to fail to meet deadlines. Some experts, too, feel that their time is better spent on revisionary work than on Flora writing; we understand and accept this. In such cases we can sometimes arrange that a contributor work in close collaboration with the expert in preparing the text.

So far, however, most people invited to contribute have accepted. At this initial stage the requirements of format and schedule are discussed and agreed. Putting the arrangement into practice can be quite a different matter. We have the full range from punctual, accurate contributors to those who must be variously coaxed, threatened, and pressured into producing a manuscript.

If there is no expert, we check around the herbaria to see if a staff member has an interest in the group and is willing to take it on. If this still fails then we have the alternatives to write it in-house or to advertise through our grants program. The latter course is preferred for any group larger than, say, 20 species.

The ABRS grants program is designed to support research in little-known taxa (both flora and fauna) and attract contributors to the publications. Currently some A$800,000 is available annually, and is split roughly 50-50 between flora and fauna studies. The taxonomic goals are controlled by setting what are called Preferred Objectives that are listed in the advertisement calling for applications. The Objectives are determined by consultation with the Editorial Committee and the Advisory Committee and are related to the publication program. Thus, for example, if we have an unallocated family for a volume due for completion in five years' time, it could be advertised to attract a contributor.

Applications compete on the basis of taxonomic priority, competence of applicants, time scale, and justification of funds requested. There are always more applications than can be funded. Funds provided through grants range from salary and travel expenses to assistance for technical or field work.

A useful planning exercise has been a series of workshops, held to discuss the preparation of the larger families and the non-vascular groups. Each lasts one day and discusses the state of knowledge, problem areas and what is needed to resolve these, the overall classification, and a plan for writing the Flora account. We find 6-8 the best number of participants. They are selected for their taxonomic expertise and overall knowledge of the group and its researchers.

Management

Management of volume progress is assisted by the use of flow charts. On these the progress of each contribution may be seen at a glance. These have proved useful not only to the editorial staff (there are currently some 25 volumes in preparation), but also to administrative staff and management.

To assist contributors and artists we have published two Guides. The Guide for Contributors explains the format and cites examples from published volumes. We tried to find a balance, without excessive detail (which would deter users from reading it) but sufficiently clear to cover most requirements. The value of a good index has been shown in this Guide.

Despite the Guide and the existence of eight published volumes of the Flora, poor format in manuscripts received has been a major factor in our slow rate of production. We ask contributors to read the Guide and to submit a few sample pages at the beginning. This procedure pays off in some but not all cases. Once we have been through it, however, it is easier to return a poor manuscript for correction by the contributor, rather than have the Editorial Assistant spend time doing the work.

A further recent publication, prepared in response to a continual problem of interpretation of terms, is *Plant Indumentum: A Handbook of Terminology* (Hewson, 1988). Written by our Flora Writer but widely circulated for comment before completion, this reviews a large number of terms and provides definitions of those recommended for use in the Flora of Australia.

The Flora is computerized to the extent that manuscripts are entered on disk and all subsequent editing is by word-processor; typesetting is done through the CSIRO computer facility (CSIRONET), the commands being entered in-house, and microfilm and final bromides produced by CSIRONET; and programs have been set up to extract map captions and the preliminary index. The system works well for editing, although there are occasions when a buildup of manuscripts causes a bottleneck in keyboard work.

Typesetting in-house gives complete control over the text up to the final proofs, but has its problems. The Bureau has a small computing staff but none trained in typesetting. The III COMP 80 typesetting system is used for the Flora, and its adaption to the required format has been developed through close consultation between the editorial staff, Bureau computing staff, AGPS design staff, and CSIRONET. While the basic procedures are now well established, problems continue to arise that require outside assistance or trial and error to solve.

Keyboard entry of typesetting commands is carried out by Flora Section staff. While this is the most efficient way given the decision to typeset in-house, it does divert resources from other editing tasks. Currently a 500-page volume takes 4--6 person-weeks to typeset.

It is possible that we will shortly lose the CSIRONET typesetting facility and so we are reviewing the options of either another in-house program or sending manuscripts for outside typesetting.

Marketing

Success in producing a Flora can be judged essentially by two criteria--sales and usage. We believe we are succeeding in both.

The Flora is published and marketed by AGPS. Marketing has been a problem, in that AGPS had no experience with this kind of publication. Despite our own lack of expertise in this area, we have supplied a great deal of information on likely markets, advertising avenues, and places for review. Many overseas purchasers have had difficulty in acquiring the Flora, but the problems are gradually being resolved.

In February 1988, total sales of the eight volumes so far published stood at over 22,000. Volume 1, with general introducing chapters and a key to families and therefore of wide interest to students, has sold over 6,500 copies. The descriptive volumes appear to be levelling out at around 2,500 to 3,000 copies. This appears to be a satisfactory figure, given, for example, our population of ca. 16 million and the fact that *Flora Europaea* has sold just over 5,000 copies.

In order to make the Flora more widely affordable, it is issued in both hard and soft cover, the latter costing A$10--15 less than the hard cover version. A concern for the future, however, is the cost to new buyers purchasing a complete set.

PROBLEMS

Had we not had an expectancy of publishing three volumes of the Flora per year, our production of nine volumes (just over 3,000 pages, 70.3 families, 463 genera, 2,631 species) in 7.5 years might appear satisfactory. Although most people in Australia understand the basis for the slow progress, it is worth summarizing the reasons here. Some are, of course, dependent on additional resources, but others can be met by present contributors and editorial staff.

Format

Manuscripts submitted in poor format are the major reason for slow progress, given other present resources. For the first few years, contributors could be excused for this, but since published volumes and the Guide for Contributors have been available there is no excuse. From discussions with other editors, I understand that this is a widespread problem. We have found that correction of format can add up to 20% to editing time, a significant amount in a large work. We now ask each contributor of a manuscript of 10 or more taxa to submit a sample very early. This is checked and returned to the author. If the final manuscript is still poor, it is returned to the author for correction. With a small editorial staff we cannot afford to spend time on such work. Because the Flora is a Government project we must be seen to be making progress, in the form of solid books.

Deadlines

As with format we have the full range, from those (how wonderful they are!) who are ahead of or on schedule, to those who must be variously hounded, cajoled, bribed, or threatened to produce a manuscript. There are many reasons for this

failure. Common ones are over-commitment, lack of self-discipline, inadequate planning, and lack of motivation. Some learn from experience, others never do.

I believe, incidentally, that every undergraduate course should include a segment--even if only one lecture--on the need for self-discipline and attention to detail when writing up research. It is an aspect that many workers do not anticipate and often deters them from completing papers. How often has it been said that research is useless until the results are published.

Study of Herbarium Material

The increased taxonomic activity in the Australian flora has placed great demands on the loan service provided by herbaria. At the beginning of our Flora project there was inevitably a period when loans were needed urgently. With planning, this has now largely disappeared, but herbaria still must handle many loans, placing a heavy load on small staffs. Because most taxonomists are aware of this fact, there has been some reluctance to borrow material from all herbaria. The shortfall can sometimes be met by visiting other institutions. Usually, however, the result is inadequate study of available material, which carries across to the comprehensiveness of descriptions and especially distributional data. In one case a referee's report showed a serious problem in a large manuscript, which had to be addressed by a loan at short notice and a delay in completing the Flora volume. Herbaria whose collections are not studied have also the problem of unannotated material that their own staff must then curate.

Contrary Contributors

Delays can be caused by contributors who believe that they can improve on the format, or that their group is so important that it requires a much more detailed treatment (e.g., longer descriptions), or who disagree with editorial procedures. In these cases it is often very useful to have the work of the Editorial Committee as a backstop, pointing out that the Flora is not a personal project but a Government one planned by consensus; we in the Bureau are simply doing our best to bring it about with a high level of consistency.

With a limited number of potential contributors in Australia, it is important to avoid antagonism. Achieving this while insisting on an appropriate contribution can be a delicate balancing act.

Competing Floras

Many contributors to the Flora of Australia are also asked to write for regional or overseas Floras. While this is perfectly reasonable, it can lead to clashes of priority. In some instances we have arranged a contribution that has subsequently been late; inquiry has revealed that a later request for another work has been accepted and written before our contribution.

Complacency

The fight to convince the Government to support the new Flora was a long one. The project itself will also take many years to complete. Enthusiasm was high in

the beginning but there is a danger that this will wane to the point where contributors lose interest. There are several reasons for this. Young taxonomists trained in sophisticated methods sometimes find it hard to produce the classical taxonomic account (losing sight of the fact that this is what most users need). People see the volumes appearing and assume that "others" are producing the work. As senior taxonomists move on, and as more regional Floras are published, the new generation may feel there is less need for a national work.

Fewer Taxonomists?

Tertiary institutions in Australia are running down the taxonomic segments of degree courses. At some stage it may become impossible to recruit sufficient new people, even though we see a need for many years ahead, not only for the Flora, but in broader terms as we attempt to fully document the biota.

Insufficient Artists

Illustrations are an important feature of semi-popular Floras. In Australia there are relatively few artists competent and available for botanical illustration. Only one technical college course in natural history art is run, in New South Wales. Artists generally must develop expertise by following published works and by guidance from taxonomists. Only four Australian herbaria have a staff artist. Some taxonomists prepare their own illustrations. Others depend on freelance artists, usually contracted for a particular job.

At the Bureau we maintain a register of all scientists and illustrators who have an interest in ABRS programs. The list is used when we need to find an artist, either locally or in one of the States, to prepare illustrations in conjunction with a contributor. This is dependent, however, on availability of funds and artists.

Botanical artists, like others, tend to be unreliable for meeting deadlines.

Overplanning

It is important that the editorial and administrative structure be adequate but not topheavy. A great deal of careful and recorded planning is needed in such a complex project that involves dozens of volumes over many years. It must allow ready assessment of progress, make provision for contingencies, and cater for change of personnel. If the administrative end becomes too large, however, much time can be lost, e.g., to duplicating information for others to use at meetings once or twice removed from the workplace. The Editor should have charge of the project but have strong administrative support, thereby allowing time for final editing checks on volumes.

It is preferable that the central staff be recruited young enough and given incentive and encouragement to stay, to assist continuity. The ideal person combines taxonomic awareness (of both plants and people) with editorial and administrative skills.

PERTINENT LITERATURE

Australian Biological Resources Study Annual Report 1986-87. AGPS, Canberra, 1987.

Flora of Australia, Vol. 1, Introduction. AGPS, Canberra, 1981.

Flora of Australia, Guide for Contributors, 2nd ed. Bureau of Flora and Fauna, Canberra, 1987.

Flora of Australia, Guide for Illustrators. Bureau of Flora and Fauna, Canberra, 1986.

Hewson, H. J. 1988. Plant Indumentum, A Handbook of Terminology. Canberra: Australian Government Printing Office.

OTHER FLORISTIC AND RELATED PROJECTS OF VARIOUS SCOPES[1]

Pennsylvania Atlas and Annotated Checklist

-- Ann Rhoads

The Morris Arboretum is carrying out a computer-based Pennsylvania Atlas and Annotated Checklist project in collaboration with institutions elsewhere in the state. The initial database was a card file of ca. 250,000 site-specific collection records on dot maps by taxon, assembled by Dr. Edgar Wherry. This card file was the basis for the Atlas of the Pennsylvania Flora that was published in 1979, which reproduced the cards directly. For the current project the records have been entered into a computer database, which runs under dBASE III in an IBM PC-AT with a 72-megabyte hard disk, and has three main components:

- The Taxon File, which includes data on nomenclature, common names, native/introduced status, geographic and physiographic boundaries, weedy status, poisonous/non-poisonous status, and growth form

- The Location File, which is based on the Wherry card-file records and is being augmented by data captured from many additional specimens not recorded on the Wherry cards

- The Herbarium File, which contains data used for building specimen inventories and for generating herbarium labels

Templates have been developed for entering and editing data for the files, and various output formats are possible. The database can be manipulated to produce many useful products. For instance, checklists of species by county or by any designated rectangular area can be drawn from the Taxon File. The Location File can be used to generate distribution maps and, together with the Taxon File, to provide lists of species occurring in particular counties or specific locations. This is an evolving system, with new features and uses being added virtually every week. Together with experts in computing and landscape architecture at the University of Pennsylvania, the Morris Arboretum is now exploring how to use the database for sophisticated GIS mapping, in which the high-resolution distributional information for species can be linked with data sets for other environmental parameters such as geology, climate, and land use. Among many other benefits, such mapping provides a predictive tool for work on rare and endangered species.

The first major product of the computerized database will be an annotated checklist of the state's flora. Meanwhile, a series of keys to Pennsylvania taxa is also in preparation. Over the longer range, a manual of the flora and, eventually, a comprehensive Flora are planned.

---

[1]Reported and Compiled by John G. Packer and Robert W. Kiger

Illinois Natural History Survey: Fish and Wildlife Mapping

-- Lorin Nevling

The Illinois Natural History Survey is one of about 10 such state surveys in the United States. Illinois also has two other state surveys, Water and Geological, both part of the Energy and Natural Resources Department. The Natural History Survey operates on a total annual budget of $3.5 million and it comprises several programmatic subgroups: Fauna, Aquatic, Botany and Plant Pathology, Wildlife Services, and Economic Entomology.

Illinois is ecologically very diverse, and the state's flora is well documented by specimens, literature, and much unpublished data. The Survey has established that about 3,000 species of vascular plants occur in the state, about 900 of which have been introduced. The progressive impact of man on the Illinois forest and prairie has been mapped using historical records that have been computerized. In recent years the state's natural areas, and thus the habitat for maintaining diversity, have been very much reduced, and the Survey is working to preserve some of what remains. A recent undertaking was a survey of rare plants known to have been present in the state historically. Of the 88 species of interest, a total of 37 were relocated, but the remainder, a significant number, were not.

The Survey's mission is to fill the state's information needs relating to natural history, and those determine its program. In order to pursue this work more efficiently and effectively, a decision was made recently to computerize the Survey's data operations. The resulting networked relational database system is called PRIME, for which the development cost has been $2.5 million per year. Examples of the botanical data that can be supplied by PRIME include the natural community preferences of rare or endangered species, collections of specimens by decade, distribution of given species by county, and checklists for counties or any other subdivisions. PRIME functions within a statewide GIS scheme, which allows high-resolution mapping of floristic and faunistic data relative to many other environmental factors. It also enables predictive spatial and temporal modelling, for example to produce an optimized plan for searching out populations of rare species, based on preferred habitats.

The Illinois Plant Resources Center uses PRIME to operate a Plant Information Network as a general public service, which gets many calls and provides a variety of information on the classification, identification, threatened/endangered/rare status, biology, ecology, and distribution of plant species.

Jepson's Manual Revision: A User-Oriented Publication

-- James Hickman and Thomas Duncan

A major project currently underway at the Jepson Herbarium of the University of California, Berkeley, is preparing a revision of Willis Linn Jepson's *Manual of Flowering Plants of California*, which covers the vascular plants of California. Although the effort

is characterized as a "revision," effectively the result will be a completely new Flora not modelled on the old one. In general, the project is very similar to the Flora of North America scheme. About 8,000 species, more than one-third of the total in North America, occur in California. California is very diverse ecologically and its flora, which is notably rich for a single state, still presents many taxonomic problems. There have been three previous Floras for the state, all of which are outdated; the lack of a modern, up-to-date treatment of the plants of this diverse area prompted the development of the current project.

The Jepson's Manual project is operating on a 10-year schedule, with four and a half years remaining before planned publication in 1992. The new manual will be a single volume of about 1,500 pages, with half the species illustrated. The taxonomic treatments will all be new, and will present, in addition to the sorts of data traditionally included in manuals, information on the horticultural value of native plants, toxicity, weediness, and rare/threatened/endangered status. The manual will be oriented toward non-professional users, especially students and amateur botanists. There are 10,000 of the latter in California, and they are supporting the project fully. The treatments are being prepared by 150 contributing authors, requiring complex, large-scale coordination, as will the Flora of North America project.

The highest priorities in preparing the new manual are to: achieve the highest possible taxonomic quality; make the treatments concise, consistent, and comparable; prepare keys that work; and use vocabulary that, while not sacrificing precision and accuracy, is simple and obvious. Just past the halfway point now, the project is going well. One third of the text is in hand and the rest is due within the year. An intense job of editing lies ahead, however, to meet the 1992 publication date.

Computers are being used extensively in this project, although only six years ago there was no computer in the Herbarium. Since then, assisted by a grant from the Rand Corporation, the project has achieved full capability for computerized manuscript and book-text production. This has been the first priority in the computing efforts thus far, to help get the manual produced on schedule. Contributors now can submit treatments in electronic form as an alternative to traditional hard copy.

Meanwhile, the project has also been developing capabilities for computer databasing. With respect to California floristics, the goals in this area of computing are to:

- Provide identification aids, using the MEKA program developed at the Herbarium. Examples of such aids already prepared are a multiple-entry key to flowering-plant families, and a similar one for genera of Saxifragaceae.

- Work toward an electronic form of the new manual.

- Link glossaries and illustrations with floristic data.

The project is waiting to complete the database design and construction until all the details of book design and production for the new manual have been settled. After that, the longer-range challenge will be not only to keep the database up to date but also to find additional ways to exploit it.

The Nature Conservancy: Conservation of Flora

-- Larry Morse

A significant portion of The Nature Conservancy's program relates to the flora of the United States. Stated most generally the Conservancy's mission is to help preserve the elements of natural diversity in the United States. Three basic steps are involved: identifying places for conservation, based on identifying and locating qualifying organisms; protecting these places initially by purchase or other means; and then ongoing stewardship for preservation. The Conservancy works in close contact with the Natural Heritage Program in each state [with the exception of Alaska, which is only now developing its State Natural Heritage Program (SNHP)], and it also has close collaborative relationships with the Conservation Data Centers in Latin America and the conservation database that has been established in Canada.

A SNHP typically deals with 10--20% of the species in its vascular flora, these being the ones deemed at risk to some degree and thus needing protection. Pursuing this work involves categorizing and ranking organisms of concern or interest, analyzing relevant information, and attending to the legal aspects of protection and stewardship. A basic reference is the United States Geological Survey topographic maps, which are used in the field to plot localities in detail. Over the years data have been compiled in an extensive system of manual files, and more recently a computer system for handling all this information has been implemented. The Conservancy and the SNHPs obtain data on plants of interest from a variety of sources, including: Floras and taxonomic monographs, herbarium specimen labels, interviews with knowledgeable people, field work, and native plant societies.

At the national level the Conservancy maintains close liaison with the Fish and Wildlife Service and other relevant federal agencies, making recommendations about federal listings of threatened and endangered species. The Conservancy has also been collaborating with Dr. Kartesz and using his checklist as a standard reference, with cross-reference to state Floras as required. Information in Flora of North America volumes will be adopted as the Conservancy's standard as they are published. Liaison is also maintained with the Center for Plant Conservation, and with the United States Forest Service, which operates a database on conservationally "sensitive" plant taxa.

Environmental reviews are an important application of Heritage Program results, and the Conservancy determines its management practices accordingly. Nationwide, it now manages 3,000 sensitive plant species in its preserves. In this work it uses the G1-G5 scale that has been generally adopted for prioritizing sensitive taxa, including subspecies and hybrids, at the global, national, and state/province levels. However, there remain many taxonomic questions regarding such plants, for which Flora of North America may well provide answers. For example, whether a particular *Betula* hybrid needs protection may be uncertain because its taxonomic status is not clear and its distribution poorly known. Altogether, some 1,000 species that the Conservancy deals with are of taxonomically doubtful status, and it cannot resolve these problems itself, since its mission is not taxonomy but conservation.

The Conservancy is planning a series of Plant Characterization Abstracts, treating species of interest because of their economic importance and/or need for protection. These Abstracts will not be monographic or floristic in nature, but rather simple natural-history characterizations. In this connection, it will be very helpful if the Flora

of North America database includes information on native/exotic status, habit, economic importance, wild-harvested or not, and horticultural potential. For each species there should also be a short summary of data, including diagnostic characteristics for identification.

Floristic Checklists and Distribution

-- John Kartesz

Fifteen years ago Dr. Kartesz began an effort to compile a concordance of the nomenclature used in the many diverse works treating North American flora, and to link this with geography at the state/province level, with common names, and with photographic illustrations of the plants. Although this was expected to take two or three years, it was eight years before it was completed and a checklist published in 1980. A second edition will be published later. The checklist is drawn from a database that includes:

- Species names linked to geographical distribution
- Morphological data expressed in simple terminology, to facilitate use by non-taxonomists
- 320,000 common names
- References to 100,000 photographic slides of North American species
- 250,000 records of species distributions at the county level, which can be used to generate dot distribution maps

The morphological information is the core of the database, and a polyclave (multiple entry key) facility allows retrieval of data on species by any one or more common attributes. There is also the capability to generate checklists of taxa for any region, with resolution to the county level.

Dr. Kartesz is now cooperating with the Soil Conservation Service to merge his checklist with theirs, and the result will be published soon. He is also collaborating with the Nature Conservancy to link nomenclatural and geographical information from his database to produce checklists and distribution maps for individual states. A further step will be linking the morphological data and the photographic illustrations with the nomenclatural and geographical data, in a polyclave context such as MEKA. Meanwhile, the information in the database will be available to the Flora of North America project.

Specimen-Based Inventories Used by Zoologists

-- Barry Chernoff

A relational database system called CU-MUSE has been developed for specimen-level information on the fish collection at the Academy of Natural Sciences in Philadelphia. Data can be entered and retrieved either by individual specimen or by lot (multiple sample). The fundamental data field is the catalogue number of the specimen or lot. For each entity, the database also includes species name, collection locality, taxonomic history, and synonymy. The system has been designed to integrate both collection-management and reference functions, including distributional data for particular species, or records of associated species.

The system is relatively simple to operate and no codes are used. So far the database contains 84,000 records, which took 18 man-months to enter. At that rate it will take five more man-years to enter records for the remainder of the collections.

DISTRIBUTION INFORMATION FOR MANAGEMENT[1]

-- Warren Brigham

Experience with decision-makers in the field of planning shows that they want to have information readily available, but that they will make decisions with or without it. Sometimes they will ask for data, but when they do, the deadlines are usually unreasonable, precluding much searching for the relevant information. The larger the project, the less detail they will want. The net effect of all this is that for decisions to be based on the best available information, what is needed is a means of having all the relevant data ready to hand, flexibly permutable, and easily generalizable. Information is not useful if it is not accessible, and automation enhances both speed of delivery and ease of manipulation, and makes possible some things that previously could not have been done at all. For land-use planning and management, a computer-based Geographic Information System (GIS) is a revolutionary answer to this need.

GIS consists of hardware, software, and spatial data sets that can be manipulated and combined in almost unlimited ways. Depending upon the parameters involved, the latter will contain point data, linear data, or polygonal data. One of the most useful features of GIS is that extracted data can be displayed graphically in very precise and revealing overlay maps. A CAD (computer-aided design) system is not a substitute for GIS, because it is inherently "ignorant" (i.e., it lacks the capability to manipulate spatial data sets "intelligently"). Generally, it seems to be difficult to get funding for GIS use in biology, but the temptation to opt for something cheaper should be resisted, because for the purpose that would be a waste of money.

The Illinois Natural History Survey has used GIS in plotting the historical pattern of forest cover in the state, to assist land-use planning. The results show that in 1820 the native forest covered 39% of the state, versus only 12% today, and that its species composition has changed appreciably, with dominance shifting from economically desirable hardwoods like oaks and hickory to undesirable maples and others.

GIS has also been used in an investigation of the marsh rice rat in Illinois. For this rodent there were only 22 records in Illinois, which suggested that it could be an endangered species. If so, a management program would be needed. Maps of its distribution in the state were readily available from the database, but more information was required, such as what determines the distribution in Illinois. Data relevant to this question (habitat, associated species, etc.) could be retrieved at various levels of resolution, such as 200-acre squares or 2/3-acre squares (about 100 per square mile). At the lower levels of resolution it was easy to compare species distribution with various natural features included in the database, such as the glacial boundary, physiographic divisions, and ecological divisions. However, none of these showed any significant correlation with the rat's distribution. With high-resolution data, there emerged a partial correlation with a particular habitat,

---

[1]Reported and Compiled by John G. Packer and Robert W. Kiger

the boundary between permanent palustrine wetlands and emergent vegetation. Adding the data for another similar habitat, bottomlands, did not improve the correlation.

Eventually it became apparent that the collection sites were correlated not with any natural ecological feature but with highways, except for one site. However, when data for railroads were added, this record was found to be associated with a right-of-way. So, the habitat for the marsh rice rat in Illinois was finally determined to be disturbed marshy areas adjacent to highway or railroad rights-of-way, where construction had brought about local changes in drainage systems, causing water accumulation to form habitats for species like *Typha latifolia*. These were exploited by the marsh rice rat, and at the same time were an easy avenue for dispersal.

A study of aquatic beetles in the eastern U.S. also benefitted from GIS. This involved two species whose ranges are geographically distinct, except for slight overlap in eastern Illinois. From specimen collection data and use of the GIS database, including information (70 parameters) on such things as percentage of forest cover, temperature (all aspects), and precipitation (all aspects), it was possible to build up a detailed picture of the beetles' apparent environmental requirements. Plotting the full extent of this particular environment would then provide a hypothetical distribution. When that was done, all the collection records fell within the hypothetical range, except for one specimen from North Carolina, which ultimately turned out to be misidentified. It is interesting that, given a sufficiently sophisticated database, one can derive distribution from habitat, or habitat from distribution. This particular case also demonstrates another advantage of GIS for those studying populations of species at the edges of their ranges, e.g., for research on evolution and population genetics. It provides a ready way to know where the fringe habitats are, so that collecting and field studies can be targeted precisely.

Looking to the future, we can see clearly that hardware is getting faster and cheaper, and its potential is increasing rapidly. It is also getting smaller, which is very significant considering that space is always critical. Networking is becoming increasingly important, including between databases. Data themselves are becoming not only more plentiful but better in quality, and we should bear in mind that for any particular application better data may exist in some format or context other than our own, and may be accessible through networking. Finally, although software is also improving, far more creative applications are possible and should be promoted. In particular, we should concentrate on better presentation of data to the user, and thus encourage broader and more effective use of the ever-growing body of biologically relevant data.

FLORA OF NORTH AMERICA PROJECT[1]

-- **Nancy R. Morin, Theodore M. Barkley, Leila M. Shultz, Robert W. Kiger, and Marshall R. Crosby**

The current Flora of North America project developed after a meeting of botanists held at the Missouri Botanical Garden in 1982 to discuss the subject. Early interest in writing such a Flora is described here briefly. *Flora North America*, really the direct predecessor of the Flora of North America project, is dealt with in some detail in part because there has been confusion about the sequence of events and in part because the new project has adopted a great many of the ideas and approaches used by the old. The Flora of North America editorial committee was assembled in 1984, and the organization of the project has developed as a result of deliberations of that group over the past four years. Both the editorial committee and advisory panel reviewed major aspects of the project as they were presented in a funding proposal written in 1985, and the project has been refined and modified in response to new ideas, changing technology, and suggestions made by people within and outside of the project. Thus, the plan outlined here represents the thinking of many North American botanists. We begin with what we think is a very solid base, and perhaps because of that we come to this workshop open-minded, eager for new ideas and suggestions about how the project can best serve the needs of the many users of floristic information. The project has enough strength and support already to allow it to expand or modify in response to the ideas that emerge from this workshop.

Historical Overview

Botanists have recognized the need for a Flora of North America for more than 175 years. Frederick Pursh's *Flora Americae Septentrionalis*, published in 1814, was the first work attempting to treat the entire area, and in 1818 Thomas Nuttall published his *Genera of North American Plants*. John Torrey and Asa Gray began publishing a *Flora of North America* in 1838, but only six of the seven parts were completed. Gray continued later with a *Synoptical Flora of North America*, published in 1875 to 1897, but this work was not completed either. The ongoing, monographic series *North American Flora* was begun in 1905 by Nathaniel Lord Britton and produced by the New York Botanical Garden. It consists of fascicles issued sporadically that treat the plants (including fungi and algae) of North America (including Central America), often in monographic detail. Despite a long history of interest, there still is no single modern work that treats all of the vascular plants of North America north of Mexico systematically and consistently.

---

[1] Compiled by Nancy R. Morin

Flora North America

Partly in response to the announcement at the Tenth International Botanical Congress in Edinburgh in 1964 that Volume 1 of *Flora Europaea* was soon to be published, North American botanists began to consider launching a similar project at home. At the annual meeting of the American Institute of Biological Sciences (AIBS) in August 1965, the Council of the American Society of Plant Taxonomists (ASPT) appointed a committee to study the feasibility of undertaking a Flora of North America.

This committee, chaired by Robert Thorne, concluded that production of a Flora of North America was both feasible and desirable at that time. In 1966 the Council of ASPT voted to sponsor the Flora project, to be called Flora North America, and designated an organization committee under the chairmanship of Mildred Mathias. Mathias drew up an organizational structure in which the Smithsonian Institution served as the project headquarters.

In 1967 sponsorship of the project was transferred to AIBS. The National Science Foundation awarded two grants to AIBS in July 1969 to support planning for the project. Funding was also received from the Information Systems Program of the Office of Science Information Service. The grants provided for a small central staff at the Smithsonian and for travel by the planning committees. By 1969 automated data processing had assumed major importance, and part of the planning grants funded a Systems Development Manager. Peter H. Raven, Chairman of the Editorial Committee, was Program Director, and Stanwyn G. Shetler was the Program Secretary.

The planning phase was productive. Pilot projects were developed in order to define the Flora North America databasing approach. The biological community was canvassed with several questionnaires designed to identify potential users and to inventory plant systematists and evaluate their interests and potential as contributors. Pilot databases included: The Type Specimen Register, which included 7,000 records at the end of 1971, "authority files" such as an author file, a higher taxon list, a species name list, a contributor file, and a morphological/ecological vocabulary.

The Program Council, formed in 1971 to advise AIBS, recommended that the Smithsonian Institution become the administering institution, taking that responsibility from AIBS with full approval of AIBS, on 2 July 1971. The Secretary of the Smithsonian made a commitment to the National Science Foundation that the Smithsonian would take on an increasingly greater proportion of the funding of the project and would maintain the database indefinitely.

A proposal was submitted to the National Science Foundation in 1971 requesting funds ($8,122,640 for six years) for the implementation of the project. Similar proposals were submitted to the Canada Department of Agriculture and the National Research Council of Canada. The funding, which was approved, was to support seven regional editorial centers in Canada and the United States and the headquarters and systems development at the Smithsonian Institution. After six years of planning, the project became operational in the fall of 1972. Continued Smithsonian Institution support depended on approval from Congress for increased funding. The United States Office of Management and Budget late in 1972 ruled that the Smithsonian Institution could not approach Congress for expanded

funding. Without Smithsonian participation, the National Science Foundation was compelled in early 1973 to withdraw its support as well, and the project was suspended that summer.

Between 1972 and 1982 several approaches were used in an attempt to secure funding for Flora North America. In 1978 Peter Raven, on behalf of the FNA Program Council, discussed with the National Park Service the feasibility of a joint project between the U.S. and Canadian Man and the Biosphere Programs (MAB) and FNA. James Reveal working with the New York Botanical Garden prepared a proposal for such a project, which was presented to the MAB/FNA Council. The National Park Service provided funding for the first year, during which Reveal, stationed at the University of Maryland, established a format for the Flora, selected a classification system, began contacting potential authors, and wrote an expanded proposal for continuation of the project. At the New York Botanical Garden, in cooperation with the Parks Study Unit, J. I. Lawyer and colleagues assembled a bibliography on state-level floristics of the U.S. and Canada. Through a subcontract to the Hunt Institute for Botanical Documentation a report on geographical interests of plant systematists, particularly regarding studies of National Park Service Lands, was prepared.

Unfortunately, continued funding from the Park Service was not forthcoming, nor did the Service fund a related project proposed by C. R. Bell and his colleagues. Subsequently, Bell secured private support to begin an "Atlas of the Vascular Flora of North America."

History of the Flora of North America Project

In spring, 1982, twenty-two systematic botanists from the United States and Canada met to consider a plan proposed by the Hunt Institute and the Carnegie Museum to resume collaborative preparation of a Flora of North America north of Mexico. In this new plan, a significant portion of the project costs would be contributed in kind by participating institutions, and the Flora would consist of traditional treatments prepared by members of the systematic botanical community. These treatments would be used to develop a database.

At the annual ASPT Council meeting in August 1983, the Missouri Botanical Garden offered to serve as the organizational center for the multi-institutional, collaborative Flora project and offered the time of its Administrative Curator, Nancy Morin, for preparation of a proposal and project organization. Morin developed a structure for the project, and assembled an Editorial Committee and an Advisory Panel.

Funding

Members of both the Editorial Committee and the Advisory Panel helped with the preparation of a funding proposal that was discussed at the 1984 Council and Business meetings of ASPT and submitted to the National Science Foundation in 1985. Funding for that proposal was declined, and the proposal was revised and resubmitted in 1986. The editorial committee agreed among themselves to proceed with preparation of volume one even if outside funding could not be secured. They began identifying potential authors, talking with publishers, and moving ahead in general. Although NSF did not fund the second proposal, a grant of $100,000 was

made by the J. N. Pew, Jr., Charitable Trust with an additional $100,000 to be awarded if matching funds of $200,000 could be found. The Pew grant supported administration of the project, including editorial committee meetings and publication and mailing of a Newsletter carrying information on floristic work being done in North America.

A third revision of the proposal was submitted to NSF in 1987 and the announcement that this proposal had been funded was made at the beginning of this workshop. This funding provided the match needed for the second half of the Pew grant and supported work on the first two volumes. It also funded this workshop, in part.

Community Support

In response to the interest of its membership in a "Flora of North America", G. L. Webster, ASPT president, appointed an *ad hoc* Committee on "Floristics of Temperate North America" in 1982 to provide liaison between ASPT and various floristic projects, particularly the Flora of North America Project. This *ad hoc* committee was disbanded by G. T. Prance in 1984, and in 1987 the Environmental and Public Policy committee of ASPT was charged with responsibility for filling this role. In 1982 the Canadian Botanical Association-l'Association Botanique du Canada passed a resolution reaffirming its commitment and approving the steps being taken to produce a Flora of North America. At its annual meeting in August 1983, the American Society of Plant Taxonomists passed a resolution affirming its commitment to support efforts to produce such a Flora, and in 1988 the Botanical Society of America passed a resolution supporting the current project.

Description of the Flora of North America Project

The Flora of North America project will prepare, edit, and publish a synoptic Flora of the vascular plants of North America north of Mexico: the continental United States, Canada, Greenland, and St. Pierre and Miquelon Islands; and develop a computerized database therefrom. No such work has ever been completed for this area. The Flora will synthesize available knowledge about the plants found in the region and will delineate taxa and geographical areas in need of further, more critical, study. It will serve both as a means of identifying plants within the region and as a systematic conspectus of the North American flora, not only for theoretical work in taxonomy, ecology, and phytogeography but also for practical use and general reference in biology, wildlife management, forestry, range management, other environmental sciences, horticulture, and agriculture. A primary benefit will be the stimulation of theoretical and applied research.

General Description

Flora of North America will be published in twelve volumes issued over a period of twelve years. It will include identification keys, summaries of habitat and geographic ranges, pertinent synonymies, descriptions, chromosome numbers, phenological information, and other significant biological observations. A comprehensive, consolidated bibliography and complete index will be published in the last volume. About 17,000 species will be covered (Soil Conservation Service,

1982, National List of Scientific Plant Names Volumes 1 and 2, U. S. Department of Agriculture, Washington, D.C.), approximately 7% of the world's total species of vascular plants. The treatments of species will be written and reviewed by experts from throughout the systematic botanical community and will be based on original observations of herbarium specimens supplemented by a critical review of the literature. Data contained in the treatments and additional information not appropriate for publication in the Flora will be recorded and maintained in a computerized database. Sources of all data will be documented. Information will be added to the database after the Flora itself has been completed, and the database will be maintained as a permanent resource and made available through hardcopy, magnetic, and other media.

Arrangement

Taxa will be arranged systematically. The Flora will document an enormous number of personal observations and conclusions made by botanists throughout the world. In many cases these observations and conclusions will be discussed further in short papers, revisions, or monographs (authors will be encouraged to publish these in *Systematic Botany*, *Systematic Botany Monographs*, or the *North American Flora* series, as appropriate). In a majority of cases, however, the Flora treatments will be the only reflection available of the authors' systematic conclusions. Conversely, someone consulting the Flora is likely to want to know something about the affinities of the taxa. An entomologist studying the mite parasite of a moth may want to know which plant species are thought to be most similar to or most closely related to the species eaten by the moth larvae in his study, and he may want to know how those species differ in morphology, distribution, and habitat. Such information can be abstracted easily only if the taxa are arranged systematically. If no systematic arrangement is possible for a given group, the taxa will be arranged alphabetically. Problems still to be solved or level of confidence in the arrangements can be frankly stated by the authors in brief discussions. Conservative circumscription of genera will be encouraged; the extensive review process will insure a fairly even treatment of taxa. The synoptic character of the Flora's descriptions will be counterbalanced by the inclusion of references to monographic or expanded floristic treatments.

Contents

Representative or unusual species will be illustrated, as will important diagnostic characters. In some cases, tabular comparisons of characters may be presented in addition to keys. Keys will be dichotomous and leads will be strictly parallel. Synonymies will include names used in Floras of states or larger areas. Authors must account for names accepted in about twenty-five designated Floras, in the National List of Scientific Plant Names (Soil Conservation Service, 1982, National List of Scientific Plant Names Volumes 1 and 2, U. S. Department of Agriculture, Washington, D.C.), and in Kartesz and Kartesz (1980, A synonymized checklist of the vascular flora of the United States, Canada, and Greenland. University of North Carolina Press, Chapel Hill, and revision). Additional synonyms may be included if necessary. Basionyms will be cited, and additional synonyms will be listed in the database. Types will not be cited in the published volumes, because in

many cases they are not known and their designation would require more monographic work than available time allows. Types will be listed in the database as information on them is available. Taxa for which types have not yet been determined or located can be listed by the computer, thus highlighting work that needs to be done. Vernacular names will be included. Regional coordinators and reviewers will determine which names have only marginal use and which are the most appropriate (most commonly used) names to include. Chromosome number(s), when known, will be cited at the end of the description. References for numbers cited will be stored in the database. Distributions will be illustrated by small maps and described by listing (alphabetically) states and provinces in which the taxon occurs or in which it is absent, whichever requires fewer characters. Endemics and introduced taxa will be indicated as such in the text and in the database, as will threatened or endangered status. Taxonomic problems will be discussed briefly following distribution statements. Taxa in need of more detailed study will thus be highlighted in the Flora; in addition, they will be flagged in the database.

Publication Schedule

Volume 1 will contain introductory material, including brief chapters on history of floristic botany, physiography, evolution of the flora, and phytogeography of the Flora area; and treatments of the seedless vascular plants and gymnosperms. This is scheduled for completion in 1989, with publication in 1990. Subsequent volumes will appear in the years indicated:

Volume 2, Magnoliidae and Hamamelidae, 1991
Volume 3, Caryophyllidae, 1994
Volume 4, Dilleniidae, 1995
Volume 5, Rosidae: Rosales and Fabales, 1996
Volume 6, Rosidae: remaining orders, 1997
Volume 7, Asteridae: Gentianales, Solanales, Lamiales, Callitrichales, Plantaginales, 1998
Volume 8, Asteridae: Scrophulariales, Campanulales, Rubiales, Dipsacales, Calycerales, 1999
Volume 9, Asteridae: Asterales, 2000
Volume 10, Liliopsida except Poaceae, 1992
Volume 11, Liliopsida: Poaceae, 1993

A final volume containing a consolidated index and the bibliography will also be published in 2000. Keys to families will be in each volume.

Collaboration with Grass Manual Project

The United States Department of Agriculture has provided funding for revision of the *Manual of the Grasses of the United States* (A. S. Hitchcock. 1935, second edition revised by A. Chase. 1951. USDA Miscellaneous Publication 200. United States Government Printing Office, Washington, D.C.). This standard reference for agrostologists throughout North America is considerably out of date and in need of revision. Mary Barkworth at Utah State University has begun work on a

plan to produce such a revision in collaboration with other grass taxonomists. In order to assure careful coordination between the projects, James Estes, editor for Poaceae for Flora of North America, serves as liaison between the Grass Manual project Editorial Board and the Flora of North America Editorial Committee. Most authors who have been invited to write treatments for the Grass Manual will be invited to treat the same groups for the Flora of North America. Treatments for both works can therefore be prepared in tandem. Similarly, many of the same taxonomic reviewers will be used for both projects. Manuscripts for Poaceae for Flora of North America will be reviewed following the procedures for all other treatments. Preliminary style and format decisions for the Grass Manual have been made, and it is anticipated that, with the computer programs now available, it will be comparatively easy to modify the records to suit the needs of each project. The Poaceae volume for the Flora of North America will match the other volumes of that series.

Organization of the Flora of North America Project

The foundation of the Project is the many specialists who contribute and review treatments. Authors, reviewers, and all other participants will be acknowledged fully and appropriately in the Flora. This project, and its published and computerized products, will synthesize the existing knowledge of the vascular plants of North America. To do so, it must identify and draw on the expertise of the entire systematic botanical community. The structure of the project is designed to encourage maximum participation by that community in writing and reviewing the Flora.

The governing body of the Flora of North America project is the editorial committee (Table 1). This consists of (1) taxon editors, each of whom is responsible for soliciting and editing treatments of blocks of families; (2) regional coordinators, each of whom acts as liaison between the project and the botanical community in his or her region; (3) a bibliographic editor; and (4) a convening editor, who has ultimate responsibility for the coordination, quality, and completion of the project. Special advisors with expertise in specific areas, such as nomenclature or conservation, may be asked to join the editorial committee. Each editorial committee member works at his or her own home institution. The convening editor is located at the Missouri Botanical Garden, which houses a small administrative and database support staff. The bibliographic editor is located at the Hunt Institute for Botanical Documentation. Research associates will be stationed at major herbaria and will spend part of their time filling requests for information from Flora participants. The Advisory Panel (Table 1), originally composed of chief administrators of major North American botanical institutions, was expanded to include representative botanists involved in major regional floristic projects. The Panel has advised the convening editor and the other members of the editorial committee about the organization and execution of the project from the standpoints of both the systematic botanical community at large and their own institutions.

In order to meet production deadlines, manuscript submission and review schedules will be rigorously maintained. Deadlines allow sufficient time for reassignment if a treatment is not completed by the original author. Taxon editors receive manuscripts directly from the authors. If there are gross inconsistencies or

TABLE 1. FLORA OF NORTH AMERICA EDITORIAL COMMITTEE AND ADVISORY PANEL

Editorial Committee:

George W. Argus, National Herbarium, Ottawa
Theodore M. Barkley, Kansas State University
David E. Boufford, Harvard University Herbaria
Luc Brouillet, Université de Montréal
James Estes, University of Oklahoma
John Fay, Office of Endangered Species, U.S. Fish and Wildlife Service
Ronald L. Hartman, University of Wyoming
Marshall C. Johnston, University of Texas
Robert W. Kiger, Hunt Institute for Botanical Documentation
John McNeill, Royal Botanic Garden, Edinburgh
Nancy R. Morin, Missouri Botanical Garden
David F. Murray, University of Alaska
John G. Packer, University of Alberta
James B. Phipps, University of Western Ontario
Leila M. Shultz, Utah State University
Richard W. Spellenberg, New Mexico State University
Gerald B. Straley, University of British Columbia
John L. Strother, University of California
John W. Thieret, Northern Kentucky University
Grady Webster, University of California
R. David Whetstone, Jacksonville State University

Pteridophyte Advisors to the Editorial Committee:

Alan R. Smith, University of California
W. H. Wagner, University of Michigan

Advisory panel:

Frank Almeda, California Academy of Sciences
David M. Bates, Liberty Hyde Bailey Hortorium
André Bouchard, Jardin Botanique de Montréal
William C. Burger, Field Museum of Natural History
Thomas Duncan, University of California
C. G. Gruchy, National Museums of Canada
V. L. Harms, University of Saskatchewan
James Hickman, Jepson Herbarium, University of California
Almut G. Jones, University of Illinois
Robert W. Kiger, Hunt Institute for Botanical Documentation
Mark M. Littler, National Museum of Natural History
Jim Massey, University of North Carolina
G. A. Mulligan, Biosystematics Research Institute
Gerald B. Ownbey, University of Minnesota
J. G. Packer, University of Alberta
Donald H. Pfister, Harvard University
Ghillean T. Prance, The New York Botanical Garden [now Royal Botanic Gardens, Kew]
Robert L. Shaffer, Herbarium of the University of Michigan
Benjamin Stone, Academy of Natural Sciences of Philadelphia
Robert F. Thorne, Rancho Santa Ana Botanic Garden
Robert Trottier, Biosystematics Research Centre
B. L. Turner, University of Texas

departures from accepted style, the manuscript is returned to the author for rewriting. If the manuscript is basically acceptable in this preliminary sense, a copy of it is sent to the organizational center, where it is typed or machine read into the computer. The resulting printout is duplicated and sent to the editorial committee, which makes an initial evaluation and recommends reviewers to the taxon editor. The taxon editor and one other editorial committee member are required to review the manuscript, but any of the other committee members may comment on the manuscript by a specified date. If extensive revisions are necessary, the taxon editor sends suggestions for revision to the author. If it is deemed acceptable for general review (or after the first revision is returned by the author), it is duplicated and sent to the recommended subject and geographical reviewers by the organizational center. Reviewers send their comments to the taxon editor, who sends final suggested changes to the author. Final revised manuscript approved by the taxon editor is sent to the organizational center, where staff revise the computer record, distribute fair copy to the other editors, and mark copy (or code the tape) for the printer. The convening editor supervises preparation of all manuscripts for the printer.

Flora of North America Database

The database of the Flora of North America project will make available, *for the first time*, names, full descriptive information, and distributions *prepared by specialists, rigorously reviewed*, and presented in a consistent and comparable manner. The database is entered into and maintained on computers so that *for the first time* plant characteristics can be analyzed and compared in the context of their hypothesized phylogenetic relationships and geographical and ecological distributions. The data format is structured to meet two primary goals: comparison of characteristics across all taxa, and production of detailed, precise, computer-generated descriptions. The software and data structure provide an advanced multiple-entry key function for detailed comparisons and analyses that will assist in the writing of such keys. They do not allow easy production of dichotomous keys, however.

The hardware and software used allow rapid and extensive searches, report generation, and data manipulation far beyond what is available through other programs or on other kinds of hardware. Because of the preciseness of the descriptive information, searches will need to be phrased carefully so that all permutations within a more general state are located, and programs will be developed to assist with this. The benefit gained from such precision is clearly the level of information content held in the database.

The potential for future research using the information in the computerized database is almost limitless. Phylogenetic comparisons, correlations among morphological characters, linking of phylogenetic and morphological characters to information on habitat, ecology, chemistry, genetics, etc., found in other computerized databases such as the Plant Information Network (PIN), Germplasm Resources Information Network (GRIN), and the database planned for California plants at the University of California, Berkeley, make the database a phenomenally powerful research tool.

The database management system developed at the Missouri Botanical Garden is called TROPICOS and is a series of interrelated files about vascular plants and

mosses. The programs were written by R. E. Magill in collaboration with M. R. Crosby. This system has been used for a variety of projects at the Garden, as well as for curatorial and bibliographic procedures. The hardware and software work extremely well for the main purposes of Flora of North America: managing data on names, descriptions, and distributions of plants. As the project grows and as new needs and purposes are identified, the equipment and programming will need to be modified or replaced. This workshop provides excellent insight into what some of those needs will be. It will be used as a "working document" as we plan for the coming years.

Data for each taxon are recorded and maintained in a computerized database using information from the treatments written by specialists and from separate optional data provided by the specialists. These data will be reviewed much the same way that the treatments are reviewed. The stored information will include: accepted name, synonyms, nomenclatural type, vernacular name, full description, breeding system(s), chromosome number(s), habitat, phenology, economic uses, threatened or endangered status, and distribution. The building of the database will not delay preparation of the printed Flora in any way. Authors will be assured that information beyond the scope of the Flora should be included only if it is already available; further information will be added to the database after the treatment is completed. Such additions will be reviewed in order to determine their acceptability and the appropriateness of including them in the database.

The database for the Flora of North American will also be used to store information that is not appropriate to a printed, synoptical Flora (e.g., detailed habitat description), or that is known only for some of the taxa (e.g., holotype information), as well as information received after a treatment has been published. With the information stored in a database, revisions or supplements to the Flora will be comparatively easy to prepare and will be accurate and up-to-date.

Hardware

The files are maintained on an Applied Digital Data Systems, Inc. (a subsidiary of NCR) MENTOR 5300 computer with two 300-megabyte hard disks and one megabyte of working memory. It uses 9-track, 1600/3200-bits-per-inch, phase encoded, 1/2 inch tape drives, with an American Standard Code for Information Interchange (ASCII) and EBCDIC character sets and four 1200-baud modems. The system can accommodate 2400- to 19200-baud modems as well. It has 64 ports that can be used for terminals, modems, or printers. Data can be transferred by telephone, tape, floppy disk, disk pack, microfiche, or paper. We have transferred data to a large number of different computers, including mainframes, minis, and micros, without problems.

Operating System

The release 2.2 MENTOR's operating system is a PICK-based operating environment that performs not only standard file-handling functions but is designed to enhance development of database management. It handles as operating system functions many operations that in other systems would require special application programs or add-ons. For example, sorting and listing functions are available as system functions. Unlike most other microcomputer database management

systems such as dBASE and R:Base, which run on top of the computer's operating systems such as MS-DOS, PICK is designed to use available disk storage to its maximum by not requiring a fixed field-length format. Data files occupy only the space of the data entered plus system delimiters. The operating system also allows on-line expansion or contraction of a field's data or even addition or removal of fields from records or files. The MENTOR uses a virtual memory system that expands its main memory (RAM) by assigning a portion of available hard disk space as virtual memory. This memory is directly addressable, as is RAM, resulting in an extremely powerful and fast operating system.

The ADDS MENTOR as a database management machine is designed for relational file manipulation. The PICK-based operating system provides report management software and data processing functions that use English statements for sorting or selecting data for output. Files are accessed though user-defined dictionaries and system verbs that together can produce formatted reports on a variety of media. Report production is quick and simple using dictionary pointers to requested data fields. These dictionaries provide the data format and position to the operating system and allow for great flexibility in report generation by letting the user set the parameters. Inter- and intra-file sorts are possible and selected lists of requested records can be stored for repeated use. Complicated reports can be performed on-screen or stored by a remote processor for future use. Thus programming required for most operations is limited to data input and correction tasks or special output needs. PICK in various versions runs on many computers, including McDonnell-Douglas's Microdata, DEC, Prime, and some large IBMs. In addition to mini and mainframe releases, PICK has been available for the IBM XT/AT and compatible microcomputers with at least 10 megabytes of hard disk since 1984.

The TROPICOS software can also be used with Revelation (Revelation Technologies, Inc.), which runs on top of the disk operating system (DOS) and has the same data structure as PICK. Revelation is compatible with all IBM PCs, XT/AT, PS/2, and 100% compatibles. It requires 512 K of RAM, MS-DOS or PC-DOS, and a hard disk. The IBM XT implementation of PICK and Revelation have been acquired by the Garden and the TROPICOS system is now running on an XT configured with a 20-megabyte hard disk.

Files

The files in TROPICOS now are:

(1) NAMES, the core of the databasing projects, is a purely nomenclatural file, containing Latin names and the citation of their original place of publication. NAMES makes use of the files AUTHORS, TITLES, and BIBLIOGRAPHY. There is a separate NAMES data file for each plant family, as well as a file for mosses and one for ferns. Currently, about 375,000 names of species and infraspecies are in NAMES, including the National List of Scientific Plant Names provided on tape by the Soil Conservation Service. The Garden also is obtaining, on diskette or tape, indexes of the type holdings of vascular plants from other herbaria to add to the NAMES file.

(2) GENUS INDEX, based on the *Index Nomenum Genericorum*, kindly supplied to the Garden through the Smithsonian Institution as an updated subset of the ING project, is a pointer file that places each of the 38,000 or so generic names for mosses and vascular plants in its proper subfile under NAMES.

(3) TYPES contains information concerning type specimens.

(4) SYNONYMS is a system for recording synonymy from the literature. Each recorded synonym is attributed to a literature reference, which is cataloged in BIBLIOGRAPHY. Thus it is easy to trace the source of a taxonomic or nomenclatural judgment.

(5) IPCN (*Index to Plant Chromosome Numbers*) currently exists as separate files for the four issues of this ongoing series that have been published by the Garden. Information in IPCN is tagged to NAMES. IPCN makes use of BIBLIOGRAPHY to record the citations in which chromosome numbers were published.

(6) DISTRIBUTIONS records the geographical distributions of plants at three levels of resolution, country, state, and county, or the appropriately named geographical subdivisions in that hierarchical structure. The information in DISTRIBUTIONS is currently literature-based; thus citations are recorded in BIBLIOGRAPHY.

(7) EXSICCATAE contains information found on herbarium specimen labels. It will draw upon NAMES, and information from it can be linked to DISTRIBUTIONS.

(8) BIBLIOGRAPHY is actually a series of bibliographies for particular subsets of the TROPICOS system.

(9) DESCRIPTIONS is a subsystem of TROPICOS for constructing descriptions of plants. Information from systematic treatments prepared for the Flora will be entered for each character by choosing a numbered descriptor from a menu for that character displayed on the screen. Descriptors can be modified to indicate variability, combined or listed severally, or modified by non-coded author's comments. Descriptors not listed on the menu can be simply typed in individually where needed. This program allows almost complete flexibility in the stored descriptive information without sacrificing space, speed, or data manipulation capability. Descriptive information stored in the database will be complete and fully comparable across taxa. TROPICOS DESCRIPTIONS is designed specifically for retrieval of data subsets for direct output and for description generation. It is well suited to on-line, multiple access keys and will produce reports and comparisons useful in construction of dichotomous keys. Computer generation of dichotomous keys would require partitioning the data. The data, which are detailed and as accurate as possible, can be used also for phenetic and cladistic analysis. Careful data selection and perhaps conversion to an uncoded state may be required.

In order to accommodate the character states found in all families, essentially all descriptive terms commonly used in botanical works will be arranged in glossaries and available for use. An author will be able to specify which characters and character states are required for the genus or family in question, and these will then be the ones presented in menus. For instance, the author for *Magnolia*, or the staff typist if a hardcopy

manuscript has been submitted, need not deal with characters related to compound leaves.

The software for DESCRIPTIONS is available for authors but they will not be obliged to use it. Information from printed copy provided by the authors can be entered into the database directly. Entering of manuscripts will be much faster than by traditional typing because only reference numbers are entered, and memory space is maximized because only numbers are stored. No conversion is required of the typist and the only restrictions on vocabulary used by authors are editorial ones. Cross checks are built into the system to detect errors in input. The final manuscript will differ from the original only in matters of style imposed by the editorial committee. The information will have been captured and stored in the database as the authors submitted it.

(10) AUTHORS is a file of authors of names of plants based on the Kew Draft List and supplementary information, particularly names of authors that have appeared in the literature since the Kew list was published. The entries in AUTHORS are coded in such a way that information that does not appear in the Kew list can be retrieved easily and forwarded to Kew for updating their list.

(11) TITLES contains information concerning the titles of periodicals and books pertinent to TROPICOS, now numbering about 5,000 entries. Abbreviations are included in this file, following the recommendations of *Botanico-Periodicum-Huntianum* and *Taxonomic Literature*. AUTHORS and TITLES are particularly important to NAMES, where unabbreviated versions of authors names and titles of publications are being replaced by numerical codes that correspond to the alphabetic information in these files. This assures that these names will be consistently abbreviated as they appear in various products of the TROPICOS system.

Information in the Flora of North America database is comparable to that in similar databases planned or being developed, such as the international legume database organized by F. Bisby, the European Documentation System based on *Flora Europaea*, and databases for the Flora of Australia project and the International Union for Conservation of Nature and Natural Resources. In some cases, the programs and computers are directly compatible. In others, the data can be converted easily to conform to format requirements of the receiving computer.

We anticipate that requests for information will be of several types. Specific, up-to-date information about certain taxa may be requested. Such information will be easy to extract and may be available on-line to those with access to a modem. Such use of the database will be encouraged and made as simple as possible. Requested reports that require major manipulation of data will be prepared by project staff. Use of the database will be encouraged through journal articles, workshops, and published examples and projects. Charges, if any, will be only for materials and computer time. As mentioned above, the data can be supplied on magnetic tape, floppy disks, down-loaded by modem or from the screen. Data can be provided for use in dBASE and similar programs, as a non-formatted word-processing document, or straight ASCII format that is moved up easily into most database environments.

Computer Facilities at the Hunt Institute

The Hunt Institute for Botanical Documentation will maintain a database of literature pertinent to the Flora of North America. The Hunt presently uses several IBM and Apple personal computers and printers. All equipment interfaces for communication with the Carnegie-Mellon University Computation Center mainframes and the Carnegie-Mellon University Compugraphic computer typesetter. The Hunt also has a UNIX-based workstation, giving them access to ANDREW, the network environment developed at Carnegie-Mellon. The bibliographic data will be stored at both the Missouri Botanical Garden and at the Hunt Institute for ease of cross-file reports and to increase accessibility.

Summary

Flora of North America inherits a rich history and benefits from the work done by many individuals, particularly during the past two decades. The current project is designed to make maximum use of the abilities of the many talented botanists now working on North American plants, to ensure the scientific quality of data gathered and presented, to make that information available in a timely manner, and to make optimum use of the computer technology now available.

PART II: SYNOPSIS OF GROUP RECOMMENDATIONS[1]

HOW FLORISTIC INFORMATION IS USED

Environmental Assessment and Land Management

Floras are used in environmental assessment and land management for identification of plants, accounts of names used and in what senses they are used, and, in a more general sense, as a guide to habitats or plant associations. They are used to determine the distribution of plants, whether a plant is native or not, and whether it is rare or common. Floras are consulted in the management of reserves and to determine the ground truth of satellite maps. They are used to discover possible source areas of plants for use in revegetation and for landscaping areas using native plants. Floras are used to identify ecologically sensitive sites and for risk assessment in such sites. Although conventional Floras provide only a limited basis for assessing the responses of organisms to changes in environmental conditions, they do provide relevant background information.

A series of Floras may provide glimpses of dynamic processes over space or through time. Specimen data (given as documentation in Floras) make it possible to document past occurrences and continued existence of a given population at a site, expansion or contraction of a taxon range, or the demise of a taxon in a given area.

Applied Biology

Floristic information is of critical importance to researchers, teachers, and extension personnel in agricultural and horticultural disciplines. Some of the primary areas of applied biology that routinely use floristic data are horticulture, crop development (e.g., new plant introductions, hybridization, and genetic engineering), and resource management. Although each of these disciplines has unique uses of floristic data, there is considerable overlap. Participants from all these disciplines use Floras for plant identification (whole plants and also plant parts, such as seeds), plant descriptions, and correct scientific and vernacular names. Geographical distribution, chromosome numbers, and bibliographic information are also very helpful. Many applied biologists use Floras most often to locate potential sources of plant materials when they plan collecting trips.

In addition to this basic information, other types of floristic data used by specific applied disciplines are extremely varied. Some are listed in Table 2.

---

[1] Written and edited by Dieter Wilken, David Whetstone, Kay Tomlinson, and Nancy Morin based on working group reports

TABLE 2. USES OF FLORISTIC INFORMATION IN APPLIED BIOLOGY

HORTICULTURE

- Determination of environmental range (e.g., hardiness)
- Pre-selection of taxa for optimum use in specific locations
- Estimation of horticultural value and potential uses
- Foliage characteristics (i.e., aesthetics)
- Propagation characteristics (includes those of close relatives)

CROP DEVELOPMENT

- Determining "weedy" relatives of existing crop plants
- Determining phenology of particular species (e.g., flowering and frequency of seed crops, especially in woody plants)
- Source of information about N_2 fixation in legumes
- Source of information about genetic parameters (e.g., hybridization)
- Identifying closely related species for molecular research, hybridization studies, and genetic engineering

RESOURCE MANAGEMENT

- Determining potential weediness
- Predicting potential pest and disease occurrence and spread
- Tracking distribution (spread) of weeds
- Determining natural habitats/environments of plant species
- Determining areas of high endemism
- Determining historical trends in vegetation types
- Identifying areas of potential resource conservation (e.g., use of native species in reclamation projects; water conservation)
- Determining rarity

Basic Biological Research

Biologists use Floras for identification, accurate naming, and to determine relationships and phylogeny. Floras provide information on relevant literature. They also can be used to determine the number of species in a genus, information on host-plant associates, flowering and fruiting time, degree of variability in a taxon (both genetic and environmental), and whether the plant is of economic interest. This latter may indicate the possibility of interesting chemical compounds, fibers, etc. The use of Floras by other biologists will certainly increase as the scope of available information increases.

Many fields of basic biology routinely use floristic information (see Table 3). Other fields, such as biophysics, do not presently use such information. Floristics has an impact upon most biological sciences and, in one way or another, upon all plant biological sciences. Despite the ever-increasing potential for use of such information, in most cases researchers do not recognize their dependence upon floristic information. Equally, the scope of activities into which biologists themselves are drawn is increasing.

Biology is an extremely broad and complex scientific discipline that encompasses the study of all life and its processes. Many subdisciplines of biology, such as anatomy, biochemistry, or cytology, focus primarily on a single type of data (see Table 4). In these data-oriented or analytical subdisciplines, the questions asked are those that can be answered by single types of data rather than depending upon data of many different types. Other fields of biology such as evolutionary biology, molecular biology, or systematic biology are synthetic or question-oriented subdisciplines and depend routinely on several or many different types of data. Both major classes of biological subdisciplines are users or potential users of floristic data.

Questions asked by researchers outside plant systematics are not easy to generalize. Division of research into data-oriented and question-oriented investigation, as in Table 4, serves to identify future possibilities rather than present uses. For example, anatomy and morphology are currently more commonly given in plant descriptions and used in identification than are biochemistry and biophysics, not because biophysical or biochemical data have less intrinsic usefulness but because the information is technologically more difficult to obtain. Few physiological data, for instance, have been obtained for a sufficiently broad spectrum of plants to be used in floristics. Information on cytogenetics, morphology, anatomy, and reproductive biology can be found in most Floras at present, and data from evolutionary biology and population biology are increasingly being presented in floristics works.

Phylogenetics and Biogeography

Most of the information contained in Floras is potentially useful to phylogenists. Phylogeny is the study of genealogy and evolutionary relationships among species and higher levels of taxonomic organization. Phylogenetic studies attempt to infer evolutionary history based on relative similarities of structure, genetic composition, and function, and therefore rely upon quantity and quality of data as well as a synthesis of data from as many sources as possible. Phylogenists use data from a diverse range of macroscopic and microscopic characteristics from external and

TABLE 3. KINDS OF FLORISTIC INFORMATION USED IN BASIC BIOLOGICAL RESEARCH

ACTIVITY	GENERAL NEED
Identification	- Every user needs to be able to identify
Accurate naming	- Scientific name, synonyms, authors, vernacular names, legally acceptable names
References	- Means of locating literature and knowledgeable persons
Systematic relationships	- Means of locating and discerning close relationships
Information on infraspecific taxa	- Names of species, varieties or formae into which taxa are differentiated
Information on cultivars, hybrids, escapes	- Area of origin of alien taxa
Size of group	- Number of taxa in the region, political boundary, geographic area or the world
Protection status	- Protection of taxa by afforded legal status, commerce restraints, etc.
Host-Plant Associations	- Relationships of parasitic plants to host plants
Phenology	- Times for flowering, fruiting, leaf expansion, sporulation, etc.
Economic factors	- Status as weeds, potential negative economic impact, current or useful products
Plasticity	- Environmental or genetically controlled ranges of expression
Variation	- Ranges in expression of characteristics
Distribution	- Location of documented sites of occurrence and areas of potential occurrence
Habitats	- Ecological characteristics of sites of documented occurrence

TABLE 4. PRINCIPAL SUBDISCIPLINES OF THE BASIC BIOLOGICAL SCIENCES

DATA-ORIENTED AREAS:
RELY ON A SINGLE TYPE OF DATA

- Anatomy
- Biochemistry (Phytochemistry)
- Biophysics
- Cytology
- Cytogenetics
- Embryology
- Genetics
- Morphogenesis
- Palynology

QUESTION-ORIENTED AREAS:
RELY ON SEVERAL TYPES OF DATA

- Ecology
- Evolutionary Biology
- Developmental Physiology
- Molecular Biology
- Population Biology
- Reproductive Biology
- Systematic Biology

internal structure (e.g., vegetative and reproductive structures such as leaves, flowers, life-forms, surface features of seeds and pollen, anatomical details of cell walls, stomates). Patterns of phenology including duration and season of flowering and fruiting, other aspects of reproductive biology such as types of genetic incompatibility or the presence and type of asexual systems, and precise qualitative and quantitative chemical data are also useful in determining phylogenies.

Biogeography analyzes and integrates the known distribution of organisms with all aspects of the environment, including both biotic and abiotic features. Biogeography also addresses such important questions as patterns and processes of migration, centers of endemism and origin, anthropogenic and climatic changes, and biological diversity. Like phylogeny, biogeography relies heavily upon quantity and quality of data. Biogeographical interpretations use hypothesized phylogenetic relationships (and the supporting data) to reconstruct and interpret past and present changes in the environment. Data important in biogeographical studies include physiological variation (e.g., photosynthetic pathways, limits to respiration and photosynthesis), pH requirements, tolerances to specific elements, and other aspects of substrate chemistry and structure. Information on pollinators, herbivores, seed dispersal, and abundance of associated organisms is useful in interpreting which organisms were contemporaneous. The probable occurrence of organisms in specific areas over time are studied by correlating climatic and geological data (e.g., altitude, insolation, amount and type of precipitation, origin and type of geologic formations, fossil assemblages). Chromosomal data at various levels of complexity have contributed to a better understanding of polyploidy and geographical patterns of chromosomal structure within and among related taxa.

Use by the General Public

Floristic information is of use to the general public in ways that most people never realize. Table 5 presents a list of questions that might be asked by people in a variety of walks of life, and that can be answered by floristic information. Some, such as the day hiker's need for a guidebook to the plants of his local state park, are not surprising, but others, such as a municipal planner's need to know what plants are suitable for use in a reclamation project to beautify the former town dump, or a farmer's query about what plants on his land might injure his livestock, are less obvious. The length of the list is indicative of the broad utility of floristic information.

Summary

Floras are used primarily to identify plants, to determine what names should be used, and for information on the morphological characteristics, distributions, and habitats of plants. Because Floras are usually brief accounts, much of the data to which Flora authors have access, and which would be exceedingly useful to those in basic or applied research, is not included. In other words, the information that *is* usually in Floras--characteristics, distributions, notes on natural history--is very valuable for all of the above disciplines, but Floras do not supply as much information as scientists and applied researchers would like. Most of the working groups agreed that existing Floras, despite the limits of their information, would be

used much more if (1) they were more available (or if people knew of their availability), (2) the keys worked reliably, (3) the information, especially nomenclatural information, were less abbreviated, and (4) if habitat and other ecological information were better.

TABLE 5. POTENTIAL USES OF FLORISTIC INFORMATION BY NON-TAXONOMISTS

1. ECOLOGICAL CONSULTANT
   --Full list of species and common names
   --Does species composition change with disturbance?
   --What is the species' conservation status? (rare, etc.)
   --Species substrate preference/tolerance
   --Are the names in my list reliable and recognizable by my user/client?

2. ENVIRONMENTAL ENGINEERS
   --Life form of plants
   --Life cycle time of plants
   --Tolerance of flooding (xeric, mesic, hydric preference)
   --Tolerance of atmospheric and substrate toxicity
   --Local climate

3. ENVIRONMENTAL ASSESSMENT
   --Local climate
   --Environmental sensitivity
   --Soil properties
   --Shade-tolerant species

4. VEGETATION SURVEYORS
   --What species may be expected to grow together?
   --A local key for use on a site
   --Any information from similar ecosystems

5. FOREST MANAGERS
   --Shade tolerance of seedlings
   --Animals present now and predicted to occur after proper forest use
   --Predicted species composition (naturally present) after a certain time, e.g., in 30-year-old regenerated stands
   --Rare plants represented
   --Herbicide tolerance
   --Pathogen sensitivity
   --What plants produce contact poisons?
   --What plants have fast-growing, extensive root systems?

6. SILVICULTURISTS
   --Seed germination rates and survival for naturally generated vegetation
   --Life cycle patterns of natural vegetation
   --Shade-forming capacity (i.e., deciduous vs. evergreen)
   --Leaf size of natural vegetation

(continued)

Table 5, continued.

7. MINING ENGINEERS
   --What plants can be used for erosion control
   --What will grow on my "waste tip"?
   --Will my waste kill plants growing downstream?
   --Where should I look for rare plants at the start of my project so I can site my mine?

8. RARE AND ENDANGERED SPECIES AGENTS
   --What is rare?
   --When can I hope to find a variety?
   --Does a variety prefer a particular association?
   --Are names reliable/controversial?
   --From how many locations is taxon reported?

9. FARMERS
   --Identity of pasture weeds
   --Toxic plants in pasture
   --What plants can be hedged or cropped?
   --Anything rare on my property?
   --What plants can I use to control slough bank erosion?
   --Grazing tolerance
   --Prediction of plants in field during fallow season

10. ENVIRONMENTAL LAWYERS
   --How controversial is the name?
   --What are common names?
   --Could my client be expected to know endangered status of this plant?
   --How easy is it to recognize this plant?

11. FERTILIZER MANUFACTURERS
   --Species tolerance of particular nutrients
   --Large-scale soil maps with associated species
   --Prevailing weed species in sales area

12. REAL ESTATE APPRAISERS
   --Identity of cultivated plants
   --Life history time scales for selected plants
   --Value of any native plants in cultivation
   --Weeds present

13. LAND USE MANAGERS
   --Species tolerance to grazing, trampling
   --Rare plants known or likely to be present
   --Prediction of species change in time and with use

14. MUNICIPAL AND REGIONAL PLANNERS
   --Species suitable for planting, i.e., tolerant of proposed use
   --Species identification in very local "refuge" sites
   --Plants suitable for reclamation use (on garbage dumps, etc.)
   --Weed identity, especially noxious species

(continued)

Table 5, continued.

15. AGRICULTURAL CONSULTANTS AND EXTENSION AGENTS
    --Weeds present in area
    --Correlative lists of plants, animals, and pathogens of both cultivated and weedy plants

16. WEED AND PEST CONTROLLERS
    --Plants that may actively compete
    --Natural controlling species
    --Host/food plants for pests
    --Soil preference or lack of it for weeds, valuable plants
    --Herbicide/pesticide effects

17. RECREATION MANAGERS AND PLANNERS
    --Use-tough species ("kid-proof")
    --Potentially harmful weeds
    --Visually compatible plants
    --Identification of common weeds
    --Native species suitable for planting

18. SPORTS FIELD AND GREENS KEEPERS
    --Strongly sod-forming grasses
    --Mowable species
    --Naturally low-growing grasses
    --Weed identity

19. RIGHT-OF-WAY-ENGINEERS
    --Weed species identity
    --Rare plant occurrence
    --Marker species for soil types
    --Species tolerance of any proposed management practices (e.g., fire)

20. FLOOD CONTROL ENGINEERS
    --Rare plants in locale
    --Soil water tolerance
    --Which plants will be damaged by drying or wetting caused when I change flow at a given time of year?

21. WATERWAY MANAGERS
    --What are seasonal growth patterns for weeds, e.g., water hyacinth?
    --What species are useful to support bank structures?
    --What plants in a watershed generate toxic exudates?
    --What species have fast-holding leaves (won't clog up the flow)?

22. LANDSCAPE ARCHITECTS
    --What is the mature plant habit and "form" of a species?
    --What can be transplanted?
    --What plants have attractive foliage?
    --What, if any, are the hazards of a species?

(continued)

Table 5, continued.

23. ARCHITECTS
    --How fast does a plant grow?
    --How tolerant of water stress is a plant?
    --How many and which native trees are spring-flowering?

24. FOOD QUALITY CONTROLLERS
    --What is this plant bit?
    --Are there any toxic natural chemicals in this plant?

25. POSTAL SERVICES
    --Which plants are allergenic?
    --When are they allergenic?
    --What plants are contact-irritative?
    --How can we tell if this plant/plant part is from a species whose transport is illegal?

26. CUSTOMS OFFICIALS, BORDER AND OTHER ENFORCEMENT AGENCIES
    --How can I identify illegal plants? (quick method)
    --What hazards exist from plants (contact irritants, irritants, allergens)?
    --What are legally accepted names?

27. PLANT BREEDERS
    --What is extent of character variation?
    --What taxa are in cultivation?
    --What taxa are closely related to cultivated form (e.g., same chromosome number)?
    --What characters are known to have a clear genetic basis (Mendelian-type inheritance)?
    --What are known pests and pathogens?
    --What is breeding system?
    --When is flowering time?
    --What is compatibility biology?

28. POISON CONTROL CENTERS
    --What is this plant?--with either very few bits or a verbal (emotionally deformed) account
    --What non-poisonous plants are like it?
    --Is it poisonous and how poisonous?
    --Is there a treatment?

29. PLANT PATHOLOGISTS
    --What are host plants for my pathogen?
    --Are there any related, non-susceptible hosts?
    --What are intermediate host's known relatives?
    --Are these symptoms known in other taxa (e.g., galls)?

(continued)

Table 5, continued.

30. CIVIL ENGINEERS
    --Are there species that will tell me about the soil properties?
    --What plants can I expect to find as a result of my project?
    --What species will be good soil stabilizers?

31. ANIMAL FEED COMPANIES
    --What is digestibility of shoot material?
    --Any irritating structures (thorns, trichomes, etc.)?
    --Any toxic or allergenic component?

32. SEED LABORATORIES
    --What is the correct name?
    --How do we identify or characterize seeds of new introductions?
    --Where are natural sources of seeds?
    --How do we identify seeds of wild species?

33. SEED COMPANIES
    --What is the season for seed germination?
    --What is the viability time limit?
    --What, if any, is the intraspecific variation in seed production or performance?
    --Is seed clean when produced (i.e., free of fruit material)?

34. NURSERYMEN
    --Weed seedling identity (native plant)
    --Transplant survival
    --Flowering form (size, longevity)

35. TOXICOLOGISTS
    --Is plant poisonous? What part(s)?
    --Chemical composition data
    --What are non-poisonous plants that may lead to confusion?

36. INTERIOR DESIGNERS
    --Drought tolerance
    --Low light tolerance
    --Pests
    --Contact irritation

37. DERMATOLOGISTS
    --Allergenic properties present
    --Contact properties present
    --Phototoxic properties present

38. ILLUSTRATORS
    --What are visible "key" characters?
    --What is average form, shape, color?
    --What plants would be found growing with it?
    --What are the forms of co-existing plants?
    --What animals favor it for food?
    --What animals would you not expect?

(continued)

Table 5, continued.

39. HEALTH AND INDUSTRY
    --Edible parts and inedible
    --How can I best preserve for market?
    --Any side effects of consumption?
    --Nutritive constituents
    --Potential quality-reducing agents or special handling requirements (e.g., store in light versus dark)

40. OUTDOOR ENTHUSIASTS
    --Workable keys to local flora that fit in a back pocket
    --What are the local plants with antidote value?
    --What parts of what plants are edible?

41. USERS OF NATURAL DYES, CRAFTSPEOPLE
    --What natural dye pigments occur in the flora?
    --What local plants contain dye pigments?
    --When is best collecting time?
    --What local plants have fiber or weaving potential?

42. FORENSIC SCIENTISTS
    --Identification from minimal or damaged samples
    --Toxic information on plants
    --What time of day, year is flower open?
    --Does plant rot quickly?
    --Pollen identification?

43. ELEMENTARY TEACHERS
    --Common name, local taxonomy
    --Toxic plants and antidote plants
    --Rare plants
    --Simple propagation information
    --Seed germinability
    --All I need to plan a field trip for 20 children

44. SECONDARY TEACHERS
    --List of local plants of interest and why
    --Suggestion of good local taxonomy exercises
    --Best taxa for local study
    --Poison/antidote
    --Reproductive biology
    --Good examples of the idea of "key"
    --How can I make botany interesting for unwilling students?

45. UNDERGRADUATE TEACHERS
    --What are the fundamental premises upon which the Flora of North America and Floras are built?
    --Doorways to academic floristics

46. ADULT EDUCATION
    --See #43
    --What are native plants? (both a list and definition)
    --What does the Flora of North America mean to John O. Doe?
    --My interest is _____. How can I use a Flora?

CONSIDERATIONS FOR PRODUCING THE IDEAL FLORA

Current Limitations

Data from currently available printed and electronic sources are either specific only to a particular region or are complete only for well-studied taxa. The printed Flora is the traditional means for obtaining floristic data, supplemented by time-consuming surveys of original literature when specific research directions indicate the need for revisionary work. The necessarily linear organization of printed Floras also limits the ways the included data can be accessed. No single source currently can provide data with respect to character variation and important biological or ecological attributes for all taxa in a particular region, much less for North America. Data presently available do not adequately document the full extent of species diversity, species richness, or many other aspects of biogeographical interest for major climatic, physiographic, and geographical regions. Furthermore, review of available literature often reveals lack of parallel data for all taxa and all subdivisions of a particular region. The Flora of North America project presents an immense opportunity (1) to improve the consistency and quality of many kinds of data on the plants of North America and (2) to increase dramatically accessibility of such data to a broad spectrum of potential users.

The Printed Flora

Space and static format will constrain the published volumes of the Flora of North America. Despite their inherent inflexibility, however, these volumes will provide the only possible access to Flora of North America information for most potential users, at least in the near future. Consequently, emphasis should be placed on information that is most broadly important in a wide array of disciplines and most likely to fit within the constraints of the planned format. This format allows two basic approaches to accessing information in the published volumes: (1) identification and (2) retrieval of certain taxon-based data.

All potential users need on occasion to identify unknown specimens by means of keys. For identification purposes, the traditional dichotomous key should be easy to use, including clearly worded parallel statements of pertinent diagnostic characters. When taxonomic circumstances permit, inclusion of both floral and vegetative characters will enhance versatility of such keys. For certain taxa at the family level, more than one key to genera or species may be useful, including separate floral, fruit, and vegetative keys (e.g., Apiaceae, Brassicaceae, Salicaceae) and keys to both artificial and phylogenetic groups in large families (e.g., Asteraceae, Poaceae). Special effort should be made to account for "exceptional" taxa, which differ partly from the circumscription of the next higher taxon to which they belong, and the full extent of variation in a taxon should be documented over its entire geographical range.

Content

Based on the needs common to nearly all potential user groups, the requirements for most effective use of the published text include the following:

- The name of the plant. An accurate, legitimate name and its author(s) should be supplemented with often-used synonyms and a widely accepted vernacular name when available.

- A succinct morphological description. The description should be concise but reasonably complete for all diagnostic vegetative and reproductive characters, including known, verified chromosome numbers. A commentary should be provided, when appropriate, on how a taxon differs from related taxa. Data included in the Flora must be presented in a consistent and parallel fashion over all levels of taxonomic hierarchy, so that comparisons can be made.

- Information on relationships. The arrangement of taxa, within either the keys or the text (or both), should reflect what is known about their evolutionary relationships, even if this is only the "best guess" of an expert. Such information will provide a means for approximately assessing genetic relationships and identifying close relatives.

- Flowering phenology. Inclusion of flowering phenology and habit would provide means for confirming identifications, making ecological inferences, and planning collection of specimens and living material.

- Habitat. Ecological data, including detailed habitat description and information concerning associated plant communities, would be extremely useful.

- Geographical distribution. Succinct, relatively precise descriptions of geographical distribution, including both elevation and abundance descriptors, supplemented by maps, are clearly useful in a broad spectrum of disciplines.

- Citation of pertinent literature. An abbreviated bibliography should be provided to aid library searches for more detailed literature. The Flora must be well-documented, and ideally will include an extensive bibliography of primary data sources. Users must have confidence in the accuracy of the information included.

- Illustrations. The Flora must be well-illustrated and/or provide documented sources of illustrations for each taxon.

Above all, the Flora must be easy to use. It needs to be as non-technical as possible, as long as this simplicity does not lead to ambiguities.

Editorial Responsibilities

Success in producing the printed volumes as scheduled will require careful organization, efficient management, and cooperation among all persons involved in the project.

The editorial committee, as currently structured, will ensure the direction, consistency, and accuracy required to achieve a satisfactory product within 12 years. However, the editorial committee should consider very carefully the actual workload that will arise during each phase of the Flora of North America project. Regular meetings will be needed to review and discuss evaluation and recruitment of authors, staffing, funding, and technical matters pertaining to specific treatments. Consequently, the convening editor, as chair of the committee, must give special attention to developing an agenda that provides for effective and timely progress. The convening editor will assume the responsibility of managing all committee tasks and should act as final arbiter in situations involving controversial issues.

The success of the project depends on a workforce of authors, whose responsibilities in preparing manuscripts will require adherence to formats and guidelines established by the editorial committee. However, the editorial committee should remain sufficiently flexible that unanticipated problems are effectively and judiciously solved. Decisions concerning the content and consistency of descriptions and taxonomic relationships should not compromise the important goals of identification and presentation of succinct, accurate data.

The current plan requires that each member of the editorial committee have the opportunity and responsibility to examine each manuscript. Therefore, at any given time, each editor will be reviewing manuscripts in all stages of preparation. This effort will require considerable time, in addition to each member's responsibilities to his or her home institution. Meeting publication deadlines may be facilitated by partitioning efforts in certain cases. This has already been achieved to some extent by recognizing both taxon editors and regional coordinators, whose cooperation will result in taxonomic consistency and accurate presentation of geographical information. Flora of North America should consider temporarily appointing additional persons to the editorial committee for special assignments or to help other committee members during periods of high workload. Persons with special abilities or access to specialized library and herbarium facilities should be asked to help with specific tasks such as verifying chromosome numbers or the legal status of rare taxa. Particular attention should also be given to proposals that permit additional *ad hoc* financial support for editorial staffing.

Author selection and support

The recruitment and involvement of contributing authors will be critical to completion of each volume. The selection of authors should be based partly on their taxonomic expertise with the group to be treated and partly on their willingness to prepare manuscripts within the constraints of the Flora of North America guidelines and agenda. Special efforts will be required to ensure their cooperation through all stages of manuscript development.

Funding for authors is a major concern, and several approaches were suggested. It was generally agreed that the Flora of North America project should budget funds to provide authors with subsidies or small grants for travel, summer salaries,

and sabbatical leaves. Such financial assistance is important for several reasons. Grant conferral, regardless of absolute amount, carries prestige and enhances the author's status in his or her home institution. Support for travel will enable contributors to visit large herbaria, permit the study of major collections, provide access to pertinent libraries, and thereby improve the quality of their work. Funding during summer months and sabbatical leaves, perhaps at Flora of North America regional centers, will permit some authors to concentrate their efforts on the research necessary for completion of manuscripts. Flora of North America grants will aid authors who do not have institutional support to cover expenses incurred during their involvement in the project.

Direct financial support can be augmented in other ways. The convening editor can aid recruitment and increase recognition of effort with appropriate correspondence to both authors and their administrative superiors, and by special appointments as research or herbarium associates.

Efforts leading to the printed volumes will generate new data and may require publication of detailed results, including revision of taxonomic relationships and nomenclature. Consequently the editorial committee and the American Society of Plant Taxonomists should pursue ways to encourage the publication of Flora of North America-related studies in *Systematic Botany*, *Systematic Botany Monographs*, and other journals of appropriate professional organizations. And because publications are a measure of scholarly activity and professional development, authors will benefit through participation in the Flora of North America project.

The National Science Foundation and other granting agencies should be encouraged to support research that enhances the content of Flora of North America; such support will encourage participation by the home institutions of contributing authors and will expand knowledge required to produce a consistent treatment for all taxa.

The transition from research to completed manuscripts can be assisted by close interaction between the editorial committee and authors in several ways. During the course of their work, authors may require access to the developing Flora of North America database and advice from their taxon editors. Much of this advice will be provided by a clearly written, thorough contributor's guide. Making available complete lists of species will greatly enhance the efforts of the contributing authors. The editorial committee should also consider preparing a bibliography of pertinent regional floras and checklists. Communication and interaction might be enhanced by temporarily establishing inexpensive telephonic or computer-assisted networks between the Flora of North America center and the author's home institution.

The need for additional sources of information and review of manuscripts may also require periodic involvement by knowledgeable persons at regional herbaria.

Current research leading to taxonomic revision and monographs should be viewed as separate but central and closely related to the goals of the Flora of North America project. Such research involves more detailed and comprehensive effort and requires more time than the preparation of text for Flora of North America, and published volumes of the Flora of North America should not be delayed in cases where detailed systematic studies are in progress. The primary goal of many systematists is that of basic research, and some may have no interest in providing the more concise treatments required by Flora of North America. If these researchers are unwilling or unable to provide contributions, the editorial

committee may find it advantageous to arrange for a Flora of North America staff member to work in conjunction with them to produce the necessary manuscript.

The Computerized Database

An electronic database is part of the Flora of North America and will initially include information contained in the printed volumes. However, unlike the printed volumes, the Flora of North America database represents a dynamic, expandable system that can accommodate much information useful to many biologically oriented disciplines. These disciplines will benefit from timely access to reliable and potentially complete data sets and from new and more broadly based directions of research. Although all workshop participants agreed that the published volumes will be extremely important, the potential structure, content, and use of a database system received extensive and serious attention. A number of areas of concern were identified.

Optimum use of the data in Flora of North America requires that it be consistent and parallel, that it be of high scientific quality, that networks be developed with other databases, and that steps necessary for storing precise geographical localities be taken, e.g., development of a grid system. A case can be made for the position that all information about plants and much information not about plants is pertinent to plant taxonomy, and thus the database could evolve into a center for all plant-related information. Alternatively, the primary database could be one module of a center for all plant-related information, providing the basic taxonomic/floristic component for such a center.

Because data initially will become available through treatments submitted by specialists and secondarily as additional information is gathered or as a result of networks and collaboration with other research efforts, and because of the importance of producing published volumes in a timely fashion, development of the database should take place in two phases. The primary database will be derived directly from or as a part of the preparation of manuscripts submitted for publication. This database will contain all information found in the published volumes as well as related information not suitable for publication (e.g., additional morphological information) but available at the time the manuscript is submitted. It will also contain bibliographic and related information gathered as resource material for authors. The expanded database may include specimen-based data and data on chemistry, physiology, anatomy, ecology, genetic makeup, etc.

Approaches to organizing floristic information based on either taxa or specimens were each identified as having particular value with regard to accessibility and improved use. The current Flora of North America database system is taxon-based but it can and will accommodate information other than data to be used in the printed volumes. The extent to which specimen-based data will be included in an extended Flora of North America database remains to be determined. Although specimen-based information will be limited by specimen label content, completeness, and logistic problems associated with verification, it can provide vouchered, accurately determined, and precise data that is particularly critical to finely resolved analyses and is essential to interactions with geographical information systems.

Recommendations for Database Content

Recommendations for database content and relationship to potential user groups are summarized in Table 6.

Complete nomenclature

This will be essential to ensuring reliable access to relationships among data sets, exploring for information, and for providing access to taxonomic literature.

Basionyms and their type specimens are fundamental to taxonomy and classification systems, because they provide the basis for application of names. The Flora of North America project will provide a rare opportunity to develop a register of type specimens and the herbaria in which they are deposited.

Synonyms are of broad value to nearly all disciplines, the past and current literature of which may have used names from different or incomplete sources.

Vernacular names and ethnobotanical names, when associated with scientific names in a database, will be of considerable use to the lay person, to education, and to certain disciplines other than biology. Names of cultivars and cultigens of native North American taxa are particularly important in resource management, agriculture, and horticulture.

Descriptive data

Descriptive data pertaining to all aspects and levels of plant structure and function have considerable value in nearly all disciplines.

Classification and identification rely entirely on patterns of variation in plant structure, phytochemistry, and, in some cases, on function. Consequently, inclusion of statistical ranges, means, and variances, particularly for morphological data, would considerably improve the usefulness of the database.

Knowledge of whole plant structure (e.g., life-form, habit) is as useful in landscape design and restoration as it is in community ecology.

Anatomical and cytological data also show broad use in such disciplines as forestry (e.g., comparative properties of wood), physiology (e.g., differences in leaf anatomy associated with photosynthetic systems), and crop development (e.g., chromosome number).

Physiological information may be useful in determining potential ecological distribution, agronomic or horticultural hardiness, and potential economic use (e.g., nitrifying ability).

Phytochemical data are as important to natural products chemistry as they are to understanding such ecological phenomena as allelopathy and plant-animal interactions.

Under the category of genetic information are included such information as self-compatibility, cross-compatibility with related taxa, and natural hybridization, which are important, for example, in identifying new genetic stocks for plant breeding. Molecular genetics is beginning to contribute significantly to further development of plant breeding, evolutionary biology, and phylogeny. As knowledge accumulates regarding the molecular structure and location of specific genes within the chloroplast and nuclear genomes, attention will begin to focus on the North American flora for sources of new variation. Consequently special consideration

TABLE 6. SUMMARY OF RECOMMENDED ENHANCEMENTS TO CONTENT OF THE FNA DATABASE. General categories identified by potential users are listed under "Content." Primary, potential user groups (e.g. disciplines) are identified as follows: (a) taxonomy, (b) phylogeny, (c) biogeography, (d) community ecology, (e) physiological ecology, (f) population biology, (g) reproductive biology, (h) conservation biology, (i) resource management, (j) agronomy, (k) horticulture.

CONTENT	USER GROUPS
Complete Nomenclature	all
Types	a, b, j, k
Synonyms	all
Vernacular Names	all
Cultivar Names	i-k
Ethnobotanical Names	a, c, i-k
Descriptions	all
Morphological	all
Anatomical	a, b, e, j, k
Cytological	nearly all
Physiological	all
Photosynthetic pathway	nearly all
Nitrification	d-f, i-k
Phytochemical	nearly all
Genetic	nearly all
Distributional Data	all
Geographic	all
Standardized	all
Political	all
Nativity	all
Abundance	nearly all
Historical	nearly all
Current	all
Ecological	all
Habitat	all
Edaphic	nearly all
Climatic	nearly all
Vegetation Type	nearly all
Pollination Biology	nearly all
Dispersal Biology	nearly all
Life-form	nearly all
Phenology	nearly all
Vegetative	nearly all
Floral	nearly all
Fruit and Seed	nearly all
Weediness	nearly all
Pathogens	nearly all
Herbivores	f-k
Complete Bibliography	all
Management and Legal Status	d, f, h-k
Ethnobotanical Uses	h-k
Illustrations	all

will need to be given to mechanisms for describing and including genetic data that serve as diagnostic markers for genetic strains and races.

Distributional Data

At all levels of the geographic hierarchy, distributional data are among the most powerful kinds of information with which floristic data can be compared. The published volumes will provide general views of geographical distribution for each taxon. Maximum gain in using the Flora of North America database can be achieved, however, by combining fine resolution of geographical distribution with a specimen database.

Information on origin (e.g., indigenous, adventive, naturalized) is important particularly to biogeography, resource management, and agriculture. Such information can be useful in determining, for example, the historic rate of migration of introduced species and identification of potential weeds. Abundance parameters are used most effectively, for example, to monitor fluctuations in population size of weeds and of rare, threatened, or endangered species.

Ecological Information

The broad discipline of plant ecology is often asked to address issues related to complex interactions within and between plant communities, community stability, and biological conservation. Ecological information will become increasingly critical to understanding needs for resource conservation and vegetation change. Inclusion of such data in the Flora of North America database would represent a significant investment, for example, in providing data for analyzing potential change arising from the "greenhouse effect." Furthermore, such information may actually be used in ameliorating potential change. Consequently, the Flora of North America project can offer considerable data that are valuable to ecologists and other disciplines requiring ecological data.

Such attributes as edaphic parameters, geological formation, and descriptions of plant communities or associations are critical to understanding the habitats in which specific plant taxa occur. Conversely, selected plant taxa may serve as indicators of particular substrates or communities.

One major use of habitat data is in assisting field researchers to determine the particular kinds of places in which a species might be found. For conservationists, the habitat data help determine the kinds of sites offering the best opportunities for long-term protection and management of a species. Relating habitats to life-forms or other species biology factors should also be a productive avenue of data analysis, and suggest many new research possibilities. Finally, characterization of habitats of rare species or problem exotics should help land managers and environmental planners determine which species might be affected (positively or negatively) by particular vegetation management practices or land-use changes.

The development of conservation strategies for rare plant and animal taxa requires detailed information on demography, abundance, and habitat requirements. This information helps determine the kinds of sites offering high probability for success in long term management and survival of rare plants.

Such ecological data as life form, climatic requirements, phenology, and reproductive biology are considered when assessing potential use of plants in land

reclamation. These same data contribute to understanding and predicting weediness, which in turn is valuable information in agriculture and resource management.

Information on the general biology of plant taxa includes such attributes as breeding system, pollination, dispersal of fruits and/or seeds, and phenology. Detailed knowledge of such information is critical to a number of disciplines. Agriculture, horticulture, and resource management would clearly benefit from data on pathogens, parasites, and herbivores.

Illustrative Material

Including maps, illustrations, and other graphic information in the Flora of North America database was considered important and feasible, considering the recent and anticipated advances in graphic capabilities of computer software. Detailed maps of climatic and physiographic provinces, biomes, and geographic regions with the ability to couple these to taxon distributions would aid analyses in a number of disciplines. The ability to correlate taxon or specimen distributions with geographic parameters may be achieved best by including the ability to link and interact with an independently developed geographical information system.

Emphasis also was placed on the need for both photographs and line illustrations of whole plants, their habitats, and diagnostic characters. A rapidly developing graphics technology also would permit storage and retrieval of scanning electron micrographs of diagnostic microscopic characters and colorized illustrations.

Documentation of Database Information

The minimum level of documentation required for such topics as typification and geographic distribution needs to be determined. Each contributor should indicate literature sources, herbaria consulted, type material seen, and specimens that document geographic range extensions. The degree of documentation needed must be established, e.g., whether authors should be required to cite one specimen for each unit-area of geographic distribution. All data to be directly included in Flora of North America must be verifiable, that is, documented by primary sources. Data that changes with time can be handled by periodic updates from the external source.

Specimen Databases

Specimen databases are essential to supply fine-scale, accurately determined and vouchered distribution, phenological, and habitat information. Such data will be fundamental in interaction of the database with geographical information systems, for example. In order to obtain detailed taxon distributions necessary for these purposes, the problem of computerization of extensive herbarium specimen records must be addressed. It is estimated that there are about 60,000,000 herbarium specimens in the United States (P. Holmgren, pers. comm., 1988). A number of specimen databases are already in existence, especially in smaller institutions, and a way must be found to assemble their contents, as well as specimen data to be computerized in the future, so that information can be made

available at a desired scale of resolution both geographically and at any level of the taxonomic hierarchy.

Capturing fine-grained geographical information obtained from herbarium specimens would require the development of a standardized grid system, which allows potentially powerful and useful interaction with rapidly developing geographical information systems. Although such a task will require considerable planning and effort, the potential value is enormous.

Specimen labels and related historical accounts (such as old local Floras) are the major information sources on specific instances of changes in range, particularly sites in which a species formerly occurred but cannot be found now. (The problem of showing a species did not occur at a given place in the past is much more difficult, but can sometimes be made by concluding that the area was sufficiently well known, as shown by other collections, that the plant in question would have been collected had it occurred there previously.) Availability of specimen label data in an organized form would enhance botanical research in many obvious ways, generally by increasing the awareness among taxonomic and floristic researchers of the pertinent specimen holdings of the many herbaria. Ecologists, conservation biologists, and others would also regularly consult such a resource if available. The specimen data could also be used to investigate the depth and adequacy of the collecting record for a species, or to analyze the relative degree of collecting activity in various areas at various times.

More information is needed on the function of computerized networks in order to determine whether such data should be gathered and maintained centrally or in regional repositories. Data fields need to be defined, and a consistent method for quoting distribution records is needed. Production of a gazetteer and use of computerized gazetteer data to convert conventional label data would assist greatly in the translation of traditional information to a grid or point system. It will be necessary to insure correct identifications of the specimens to be accessed by databases.

Curatorial methods will be needed to assure that computer records are aligned with specimen records, so that specimens do not have one determination in the computer and another in the herbarium, or that duplicates do not have different determinations in different herbaria and databases. The whole question of dealing with replicate specimens needs to be resolved. It will be necessary to address implications for day-to-day curatorial procedures in herbaria whose specimen databases contribute to the evolving database. Specimen-based systems tend to fail for reasons of management rather than for inadequate computerization procedures. Because practices and staff duties in a number of institutions may be affected, this issue must be handled sensitively by the database manager and committee.

Authors should be encouraged to record label information for herbarium specimens they examine in a database format using standardized fields. This information can be used to help develop specimen-databases on a taxon by taxon level. Accession of new herbarium specimens should be supplemented with the development of new kinds of informative, data-intensive labels. Herbaria with special collections, including those of unusual floristic areas and by well-known, historically active field botanists, should be solicited for such supplementary information as field notes and itineraries of collectors.

The development of a geographical information system, including accurate data basic to description of soils, geology, physiography, and climatology is needed. A standardized, continental grid system should be established with special attention given to resolving data points at fine scales of resolution and with the ability to overlay such data points on traditional political boundaries. The ability to integrate with floristic data sets and multiple entry access also should be incorporated. A geographic information system can provide essential data to a number of disciplines that rely on geographical distribution, while functioning as an important resource to its own disciplines.

Access to the Database

As the extended database is developed, high priority must be placed on ease of access and versatility in retrieving information. The system should be available and comprehensible to a broad spectrum of potential users, including those with limited computer literacy or experience. Consequently, special efforts should be made to provide clearly written user manuals, relatively simple formats for data presentation, and user-friendly software, and to keep costs to the user low. A number of specific recommendations were made to enhance retrieval and dissemination of information.

The ability to provide data either on specific taxa or on subsets of categories (e.g., abundance, life-form, pathogens). Information on taxa should be accessible either by means of any synonym or by some approach that identifies the current name used in the database. Development and use of a numerical code name that identifies the hierarchical position of each taxon may prove useful in searching for information.

From a continental Flora and associated database, it should be possible to extract floristic treatments of specific regions, including states, provinces, territories, parks, vegetation zones, etc. Specialized summaries might also be extracted (sometimes through networking with other databases), such as regional lists of poisonous plants, or lists of plants containing certain chemical compounds.

The ability to conduct searches for particular combinations of data that address specific questions. For example, a resource manager may require a list of all rare taxa proposed for protection in a particular political region, their current legal status, geological substrates, and known localities. A horticulturist may desire a list of all shrubby, evergreen, drought-tolerant taxa that can survive within a specific range of annual temperatures. A reproductive biologist may wish to test the hypothesis that monocarpic taxa (plants that flower once, then die) have different life-forms than related polycarpic taxa (plants that flower repeatedly), independent of geographical and ecological distribution. Consequently, an initial search for such information would require lists of monocarpic taxa and other species in the same genera, their geographical distribution, and data on ecological parameters that might affect growth.

The ability to conduct bibliographic searches that access selected literature on a particular taxon or category of information. No current bibliographic source specializes in topics classified by taxon or by subjects pertinent to the Flora of North America project. The Flora of North America project will afford the

opportunity, through its pool of contributing authors, to develop a bibliographic source for such information.

The ability to copy and provide information on soft disks, hard disks, and other forms of electronic storage, which could be used more effectively by individuals working at freestanding terminals. It is anticipated that a single, central database may not have sufficient staffing to conduct individual searches for particular data sets. Consequently, it may be more efficient to make available general sets of data either through remote terminals or by means of mobile storage devices that can be used with appropriate software at a researcher's home institution.

The ability to provide on-line identification, including the design of interactive, multiple entry (polyclave) keys. Identification will remain one of the most common requirements of many potential users. It is reasonable to assume that, under certain conditions, proper identification may not be successful using the printed volumes. Such conditions include the lack of complete specimens and the need to identify taxa new to the area. The ability to provide some form of electronic identification, using material at hand, may prove more efficient and will provide additional or new information in strengthening the database. Quite often, inquiries regarding identification are accompanied by requests for additional data. These kinds of requests may be handled much more efficiently through an online system that ties to a centralized database.

The ability to provide access by remote terminals via phone lines or other electronic networks.

The ability to generate hard-copy outputs that result from user- generated searches and relational sorting of data. These recommendations are clearly necessary for efficient use of the proposed database and will reduce the necessary investment in staff.

It is clear that the Flora of North America database must be "user-friendly." The user, no matter how inexpert, must be helped to present questions in a way that will deliver usable answers. The predictive value of the database should be widely publicized among the "laboratory" subdisciplines of plant biology to demonstrate the potential for experimental prediction. It is likely that the Flora of North America database will catalyze the development and production of databases in other subdisciplines of biology (e.g., phytochemistry, biochemistry and biophysics, embryology, and reproductive biology). Available information in these areas may already be sufficient so that work could start immediately to develop databases that are compatible with that of the FNA project.

Networks

Connections among databases need to be established using either existing networks such as BITNET or dedicated links between host institutions. Many databases already in existence are clearly related to the elements in the primary database, for example census, climatic, soil, and wildlife data. Others in the botanical community include herbarium information, specimen-based systems, checklists, bibliographies, authority lists, and collector lists. More distantly related are databases on plant substances, toxicity, photosynthetic rates, uses for human nutrition, and noxious weed potentials. These databases need to be identified, means of access clearly defined, and ease of access provided (see Table 7).

TABLE 7. OTHER PROJECTS AND RESOURCES (Incomplete, but intended for guidance of FNA authors and users)

FLORAS

National, international plant databases

- National Biological Survey (if established)
- The Biological Survey of Canada
- European, Taxonomic, Floristic, Biosystematic Documentation Systems
- International Legume Database and Information System
- Bureaux de la Flore et de la Faune (Paris)
- Botanical Research Institute
- Flora of China
- Flora of Australia
- Taxonomic Databases Working Group

Regional

- Rocky Mountain Region
- Vascular Plants of the Southeastern U. S.
- Generic Flora of Southeast U. S.
- Great Plains and anticipated revisions
- Intermountain
- Chihuahuan Desert Region
- So. Lake Michigan
- New England
- Britton and Brown revision
- Gleason and Cronquist revision

State/Province

State Biological or Natural History Surveys

Alabama	New Hampshire
Alaska	New York
Arizona	Nevada
California	Ohio
Colorado	Oklahoma
Florida	Pennsylvania
Illinois	Saskatchewan
Lousiana	Utah
Michigan	Vermont
Missouri	Yukon

Local

- Colorado: Plants of the Colorado Front Range
- Illinois: Greater Chicago
- Idaho: Central
- NPFlora--vascular Flora of National Park Service based on Soil Conservation Service/Smithsonian Institution
- National Park specimen-based database, ANCS, through curatorial services division National Park Service

(continued)

Table 7, continued.

Checklists--national, regional, state, local
Monographs--list should be assembled
Projects involving specialty groups
Cultivated Flora of the U. S. and adjacent Canada--Bailey Hortorium
Aquatic Plant Database--Florida
U. S. Fish and Wildlife Service Wetlands Plant List/Inventory
Seed banks
Muenscher's aquatic plants revision

SPECIMEN DATABASES

Alabama -	University of Alabama
Alaska -	University of Alaska
Illinois -	Illinois Natural History Survey
Indiana -	Notre Dame University
Kansas -	University of Kansas
Oregon -	Oregon State University
Pennsylvania -	Pennsylvania University
Utah -	Utah herbaria
Wyoming -	University of Wyoming--Rocky Mountain Region

Type collections:
Smithsonian Type Register
UC Berkeley Type Register

BIBLIOGRAPHIC REFERENCES
Taxonomic Literature II
Bailey Hortorium--seed catalogue database
Bailey Hortorium--cultivated plant database
Biological Abstracts
Gray Herbarium Index
Index Londonensis
Index Holmensis
Kew Index
Index to American Botanical Literature
Index to Plant Distribution Maps in North American Periodicals
Catalog of the Botanical Art Collection at the Hunt Institute
Australian Plant Name Index
National Agricultural Library
BIOSIS on-line computer service on botanical literature
Medline
SCI
CAB

USES OF PLANTS
Agricultural Research Center--economic plants
Great Smoky Mountains National Park
Southwest U. S. Ethnobotany (Desert Botanical Garden)
Pharmacological manufacturers association

(continued)

Table 7, continued.

Feed manufacturers association
Herbal manufacturers
USDA plant product inventories
USDA import records
APHIS interception records
FDA logs
National Institutes of Health
American Association of Nurserymen cultivated plants database
Agriculture Canada database on economic plants
Commonwealth Agricultural database
CONSERVATION INFORMATION
Natural Areas Association
Restoration Ecology
The Nature Conservancy and Natural Heritage Programs
International Union for Conservation of Nature and Natural Resources Species Survival Commission
Fish and Wildlife Service Endangered Species Program
Rare plants of Canada
Bureau of Land Management Committee on Conservation of Biological Diversity
Global Environmental Monitoring System
United Nations Environmental Programme
Center for Plant Conservation
CHEMICAL/TOXICOLOGICAL
Merck Index
U. S. Pharmacopoeia
NAPRALERT
TOXLINE--Poisonous Plant Information System
MISCELLANEOUS
Database of plant morphology (Seattle)
Weed database
WHERE TO FIND LIVING PLANTS
Botanical garden accession lists/inventories
American Horticultural Society database on plants in cultivation at botanical gardens
ABIOTIC/BIOTIC INFORMATION FOR GIS APPLICATION
United States Geological Survey
National Oceanic and Atmospheric Administration
Oakridge Tennessee Atomic Energy Laboratory
U. S. Census Bureau
Energy Research and Development Administration
National Wetlands Inventory--U. S. Fish and Wildlife Service
Historical vegetation ecology
GIS databases of selected National Parks--National Park Service
GIS databases of selected National Forest units--U. S. Forest Service

The Flora of North America database should not duplicate existing databases, but be planned and structured in such a way that networks can be set up to import and export data and facilitate correlational studies on a national and international level. It is also imperative that the database structure provide flexibility to accommodate new data fields as needs arise during and after the production of the Flora volumes.

To make exchange of information as easy as possible, data fields relevant to Flora of North America should be listed and defined, support should be provided in the form of grants or co-authored publications for data contributed by other organizations, and key data elements should be given unique codes. Data format restrictions should be minimal. The data should be readily understandable and should be in ASCII with record and field delimiters. Capability should be provided for future inclusion of abstracts, full text, and on-line graphics.

It might be appropriate to convene a workshop of representatives of organizations having relevant databases to explore links that could be established, identify gaps, explore the possibility of sharing resources, and plan further action. Many organizations already have relevant databases that will presumably continue to expand and develop. The kinds of questions to be asked include: What are the fields covered? What are the gaps? Is networking possible in this case? What are the financial and copyright implications? What are the technical problems of information exchange? What kind of corporate structure might be envisaged with participating partners contributing one or more modules? Is there a possibility of establishing a national/international clearinghouse for these databases? Similar questions have recently been addressed by several of the world's leading conservation databases and associated organizations (workshop at IUCN General Assembly, Costa Rica, February 1988) and the report of the workshop should be studied.

Database Management

A Database Policy and Strategy Committee of parallel importance to the editorial committee, and reporting to the convening editor, should be formed. A full-time database manager will be needed to implement the work.

The composition of this committee should include persons with the following expertise:

- A floristic botanist with Flora of North America ties, i.e., one who is thoroughly familiar with defined goals and implemented procedures of FNA
- A member-at-large botanist, i.e., one who represents a user or customer of the FNA product
- An expert in traditional tabular database design and management
- An expert in geographical information system use and management
- An expert in communications and networking via hard-wired lines, fiberoptics, and satellites

- An expert in hardware design and performance
- An expert in software design, and programming

The committee should be charged with the following somewhat overlapping issues for consideration:

- Understanding and helping to define the scope and objectives of the database management strategy for Flora of North America, particularly addressing what is to be included in the database and who will be the users of the database--related issues will include the elements, fields, and relations of these databases
- Determining how Flora of North America will insure the integrity, stability, and security of the Flora of North America database
- Deciding who will have access to the Flora of North America database and under what conditions; what protocols these persons will follow to read data or obtain subsets or reports; who will update, edit, and correct the data, and under what protocols
- Determining how quality of data can be assured or maximized, and what tests will be administered to data before entry to the final database
- Determining what the costs are for development, implementation, and maintenance over the term of the Flora of North America project; what hardware and software the project requires; at regular intervals, perhaps annually, what new hardware/software Flora of North America should add or convert to meet its goals and the needs of its users; what costs can be projected annually for inclusion in proposals for further funding for hardware, software, and database update

Hardware and software

Hardware should include a large multi-user environment, no restrictions in growth (or at least upward compatibility), multi-operating-system environments, and all major programming languages. General purpose equipment should be chosen whenever possible. The Flora of North America database should be developed around the idea of portability, with future system enhancements as well as data transfer in mind. Future development of components should move toward a common strategy of data structure to facilitate interaction among the various parts of the Flora of North America database, other botanical databases, and databases in other disciplines. A common data strategy will eliminate the need for most conversion programs and allow more rapid transfer of data between the various parts of the Flora of North America database.

Ancillary Publications

Completion of the printed volumes and associated database may result in the need for supplementary publications that aid or enhance their use. These may include specialized indices, an annotated bibliography to cited literature, and periodic addenda or revisions. Communication of information should not be limited to traditional printed matter, as advancing technology allows publication by means of electronic ROM discs or similar devices.

The structure of the database should be designed to permit relatively rapid compilation and publication of printed material tailored to particular user groups. Such publications may include local or regional Floras and checklists, identification manuals, identification keys that employ selected characters, and computerized multiple-entry keys. It may be beneficial to publish an abbreviated version of the Flora that is designed for field use as an identification manual. Clearly, attention should be directed to these and other needs identified during review and initial use of the printed volumes.

Public Relations

Success of the Flora of North America project also depends on public relations and education. The editorial committee and authors should make every effort to prepare volumes that are used and easily comprehended by a broad audience. A major goal of the Flora of North America staff should be to generate widespread public awareness and to foster support for the project. The mission and potential benefits of Flora of North America should be communicated particularly to educators, academic communities, political agencies, congressional delegations, public and private funding organizations, museums, conservation societies and agencies, and both amateur and professional biological organizations. A special effort should be made to identify and contact groups who have an interest in Flora of North America as potential users, including resource managers, landscape architects, commercial horticulturists, and planning agency personnel. This goal can be accomplished by personal contacts, publicity, and advertisement in the news media, popular magazines, and professional journals.

A number of other approaches may be useful in publicizing the importance of Flora of North America, some of which need involve only a minimal increase in use of current resources and staff. For instance, the mailing list for the Flora of North America newsletter could be expanded to include pertinent organizations with special interests in funding, conservation of plants, and horticulture (e.g., native plant societies, garden clubs). Outreach efforts can also be enhanced by identifying and recruiting key persons at local museums and botanical gardens who would serve as information sources and spokespersons for Flora of North America at the local level and in primary and secondary schools.

Flora of North America should consider consulting a professional publicist with experience in the area of conservation or resource management. With the addition of an attractive logo and professionally developed, travelling poster displays and audio-visual packages, Flora of North America's impact can be significantly increased.

Flora of North America also should consider organization of local or regional workshops and short classroom-oriented sessions that educate and involve interested persons directly in the Flora of North America project. Some sessions could be developed to involve participants directly in the research of contributing authors or various activities at Flora of North America regional centers.

CONCLUSIONS: NEW DIRECTIONS FOR FLORISTIC RESEARCH

Continental floristic information provides a powerful resource for answering comparative questions posed at various taxonomic and geographic levels and for using information from other disciplines. Research questions can be formulated better and applications accomplished when placed in broad context, making local awareness and activity more significant and effective. Degrees of confidence in information can be assessed, gaps determined, and "remedial" research conducted. Drawing upon other major floristic information resources such as *Flora Europaea* and the Flora of China, the North American data can help answer questions of inter- and multi-continental scope.

Taxonomic, descriptive, and identification data are essential components of a floristic information synthesis used in ecological and conservation studies. Indeed, it is the taxonomic classification that forms not only the framework or organizational structure for the published Flora and the floristic databases, but also for numerous coordinated or independent databases containing species-related information that can be consulted in conjunction with the floristic data *per se*. Providing this taxonomic classification in an information-management framework is a fundamental role of systematic biology. To the extent that the systematic community does not provide its information properly synthesized and accessible, other information and activity will develop less soundly than it should, with loss for all concerned--both people and plants.

This botanical classification also forms the major linkage between information about the plants as plants, and information about these plants and their interactions with other kinds of plants (e.g., vegetation ecology) or animals (e.g., pollinators or herbivores). A continental Flora will provide a standard system of names of plants that is essential to communication in scientific research, education, and management. The availability of a continental classification also reduces provincialism and encourages greater consistency among the classifications employed in local and regional Floras, field guides, vegetation or wildlife studies, environmental assessments, and other kinds of natural-history research that include the names of plants.

The taxonomic organization of the continental floristic information also provides the capability for determining pairs of related taxa that can be studied in comparative studies. For example, in horticulture and conservation biology, one may need to develop and practice techniques on an abundant, related surrogate before working directly with an endangered entity. Similarly, the occurrence of specific secondary metabolites, other chemical compounds, or other kinds of toxicity may be predicted in certain species on the basis of their occurrence in related taxa.

Because of its current and authoritative nature, Flora of North America (including the computerized database) will serve as a starting point for the study of systematic problems involving all North American plants. Many of these problems will be elucidated during the production of the Flora. Flora of North America will also enable the development of systematics-oriented hypotheses, and the testing of these hypotheses to the extent that the required information is available.

Furthermore, the production and publication of a Flora of continental scope will provide an assessment of the state of the taxonomic and floristic knowledge of the species, bringing the information on each up to a scientific standard, thereby giving conservationists and resource managers greater confidence that local endemics are distinct entities rather than misidentified odd variants or wide disjuncts of other taxa. Noting infraspecific variation (taxonomically recognized or not) in the floristic information will also increase awareness of the pattern of variation in a species, helping in determining which of its occurrences are most significant for conservation. If the information on variation can be combined with ecological or geographical data, many other kinds of studies will be facilitated.

Identification of individual plants is essential to most kinds of botanical research, including agriculture, horticulture, forestry, and plant conservation. While local and regional Floras and other specialized guides will continue to be used in the vast majority of routine cases, the availability of a Flora on a continental scale is important when a plant is found that cannot be readily identified by works of more limited scope. Access to this continental perspective is particularly important in distinguishing wide disjuncts from overlooked endemics or from undetected escaped exotics. Availability of the information as polyclaves (multiple-entry keys) would greatly facilitate its utility, especially for identification of incomplete or vegetative specimens, such as those often encountered in ecological studies.

Species biology and other natural-history characterizations, including phenology, form a major unit of information important to studies of distributions, habitat relations, and conservation needs. Information on phenology, for example, is important in planning field research, either in deciding when to seek a particular species, or what species may be appropriate to seek in an area at a particular time. However, little information is conveyed by generalized continental-level phenological data for widespread species; finer degrees of precision are essential. Combined with geography, ecology, or other factors, these phenological data could be used to investigate questions relating phenological patterns to other aspects of community dynamics. For instance, the return of broad-tailed hummingbirds to parts of the Rocky Mountains in spring may be predicted to coincide with the first flowering of plants with reddish, tubular flowers, but may be found to coincide with the flowering of plants with tubular flowers of various colors.

Similarly, information on breeding systems or other parameters of species biology, if collected and presented in a standard form for the continental Flora, could be used in scientific studies, for example to seek similar patterns in disparate taxonomic groups, different habitats, or different geographic regions. One interesting possibility is to seek particular patterns of life-history parameters that characterize rare or declining species. Management techniques developed for a particular species, whether a rare native or a problem exotic, might also be of interest for management of biologically similar species, regardless of their taxonomic placement or geographic location. Many of these characterization parameters would also be of interest to nature interpretation or other activities involving the general public.

Biogeographical data are one of the most important categories in a Flora, and, in a more general sense, geographical data are one of the most powerful kinds of information with which floristic knowledge can be compared. Traditionally, floristic information beyond the local level is in summarized form, usually by political

jurisdictions, natural regions, and similar polygons. Grid-cell data are sometimes used, but requires substantial effort to produce, even on a coarse scale, since the source data often do not appear that way. Detailed specimen label data provide even more powerful geographical perspectives, when used with the proper cautions regarding misidentification, collecting histories, and poorly known areas.

Geographical analysis of the floristic data, even a simple itemization of states for species, can focus research by showing gaps in reported distributions and areas in which further fieldwork would seem important. Such techniques are far more powerful when done to the county or similar level of resolution. This kind of focused field effort is particularly important in determining current distributions for rare species and for assessing present and anticipated future distributions of problem exotics. The geographical data also provide a powerful tool in initial screening of species that may be rare, declining, or endangered, although current fieldwork is then necessary to determine their actual status. Similarly, these data can be used to identify centers of endemism or disjunction, which, if not already well studied, may be appropriate priorities for further field research and collecting as well as for conservation review. These data would also assist systematists in locating field-study areas in which particular combinations of species might be found growing.

Combined with the many kinds of geographical, geological, climatological, and other data now becoming available in geographical information systems (GISs), the floristic data, when of sufficient resolution, can be analyzed in numerous powerful ways, such as examining factors that may determine distribution limits, or identifying specific places in which particular species might be sought. GIS-based research might also address biogeographic questions such as whether particular distribution patterns are environmentally or historically based (including both geological history and human history). For an aggressive exotic species, analysis using a GIS might indicate the extent of the potential range, and the areas in which it might become a management problem.

It must be noted, however, that the degree of precision needed for effective GIS work is rarely available, except for well-studied rarities. Even distribution data at a precision of, say, a quarter-degree square would require tremendous amounts of new field research for the vast majority of North American vascular plants. Nor can GIS methodology as presently available be used to assess long-term historical patterns in range changes, since the available specimen label information and other historical data can indicate only places where plants formerly did occur, and cannot document where they were *not* present.

Habitat information, and information on interactions with other organisms and with abiotic factors such as substrates or microclimate, form a fourth major category of continental floristic information. Included here are the interactions, both positive and negative, with people and with human activities. Potentially, this is one of the most powerful capabilities of the published Flora and database, but traditionally it has been one of the weakest components of floristic work. It is in this area that others seek botanical information for *their* purposes, and where the writers and editors of Floras and producers of floristic databases can communicate the wealth of botanical information deemed important. The database, even more than the published Flora, becomes more important here, and this area can be expected to grow considerably in the subsequent more enriched floristic database system.

Continental floristic data are of extreme importance in the field of conservation. Indeed, this is the area where such information may have the greatest impact. Once floristic data can be integrated with geographical information, many avenues of investigation appear. For instance, what are the overall patterns of distributions in the flora of North America? With the rampant destruction of habitat due to population pressures, what should be our land management priorities for the U.S.? For Canada? Can we document this and provide a sound basis for governmental decisions? By networking floristic databases with other conservation-oriented databases we can help our legislators establish land management responsibilities.

Using Flora of North America, it will be possible to know the status of various species, whether they are rare or endangered, whether they are native or introduced, whether they are noxious weeds, etc. Working with the database, it will be possible to discover and monitor changes in status, such as the decline or expansion of native flora, the spread of alien weedy species, and perhaps even changes in the size of populations. It may be possible to have the Flora of North America database become a repository for government documentation for endangered species so that one could go to one easily accessible source to derive specific information garnered by status surveys and rules resulting from those surveys.

Used alone, the continental floristic information that will be provided by the Flora of North America project can answer a multitude of questions concerning the plants that occur in the continental region, in this instance the vascular flora of North America north of Mexico. Combined with other taxon-oriented databases, additional research questions can be explored. Combined with other geographical data systems, patterns can be sought and biogeographical questions approached. Combined with specimen-label data networks, more detailed geographical, ecological, and historical aspects can be considered. Combined with databases on conservation status, land-management responsibilities, etc., the floristic information can be applied to natural resource problems. Finally, combined with other concerns of society, such as tourism and recreation, pride in locale, economic concerns, and so on, floristic information can be made more relevant and the importance and value of plants recognized and enhanced.

APPENDIX 1

Report of Group 1. How is floristic information used in fields such as agriculture, forestry, and horticulture? How can this use be improved or increased?

Group Members

Duvick, D. N.*
Giles, R. H.
Gunn, C. R.*
Hall, D.*
Krugman, S. L.*
Larson, G.
Leiser, A.
Nelson, C. J.*
Packer, J. G.
Rhoads, A.
Spellenberg, R. (recorder)*
Wagstaff, J.
Young, D. A. (leader)*

I. INTRODUCTION

Floristic information is of critical importance to researchers, teachers, and extension personnel in the agricultural and horticultural disciplines. Our group focused on current uses of floristic data, particularly for horticulture, crop development (e.g., new plant introductions, hybridization, and genetic engineering), and resource management.

II. FLORISTIC INFORMATION USED IN AGRICULTURE, FORESTRY, AND HORTICULTURE

There is considerable overlap between the disciplines in their requirements for floristic data, although each discipline also has its own unique applications.

A. Data and applications used by all the disciplines:

- Plant identification (plant parts [e.g., seeds] as well as the whole plant)
- Plant descriptions
- Vernacular names
- Nomenclature/standard names
- Basic information about plants (e.g., geographic distributions)
- Source of bibliographic information
- Chromosome numbers
- Source of plant materials
- Guides for planning collecting trips

---

* Asterisks indicate members of the writing team for the group

B. Applications of floristic data by specific disciplines include, but are not limited to, the following:

- o Horticulture
  - Determining environmental range (e.g., hardiness)
  - Pre-selecting taxa for optimum use in specific locations
  - Estimating horticultural value and potential uses
  - Determining characteristics (i.e., aesthetics)
  - Predicting propagation characteristics (includes those of close relatives)

- o Crop Development
  - Determining "weedy" relatives of existing crop plants
  - Determining phenology of particular species (e.g., flowering and frequency of seed crops, especially in woody plants)
  - Providing information about genetic parameters (e.g., hybridization)
  - Providing information about N_2 fixation in legumes
  - Identifying closely related species for molecular research, hybridization studies, and genetic engineering

- o Resource Management
  - Determining potential weediness
  - Predicting potential pest and disease occurrence and spread
  - Tracking distribution (spread) of weeds
  - Determining natural habitats/environments of plant species
  - Determining areas of high endemism
  - Determining historical trends in vegetation types
  - Identifying areas of potential resource conservation (e.g., use of native species in reclamation projects; water conservation)

III. ADDITIONAL FLORISTIC INFORMATION NEEDED BY AGRICULTURE, FORESTRY, AND HORTICULTURE

There was consensus in our group that for FNA to be of maximum use to agriculture, forestry, and horticulture, it must provide considerably more information than typically is included in existing Floras.

A. Additional floristic information needed for all disciplines:

- o Detailed phenological data
- o Phenotypic vs. genotypic differences
- o Notations on known ecotypic differences
- o Symbionts
- o C_4 and CAM notation
- o Ease of hybridization with close relatives
- o Crossability information
- o Chromosome numbers consistently entered
- o Pollination systems
- o Relative abundance
- o Seed dormancy

- o Chemical data
- o Molecular data (particularly regarding phylogenetic relationships)
- o Chemical relationships (in a phylogenetic context)
- o Plant/plant and plant/animal interactions
- o Habitat/community associations
- o Reproductive history
- o Specific information on sources of seed materials (e.g., seed banks)
- o Specific information on the location of living plants (e.g., in gardens/arboreta) and herbarium specimens.

B. Additional floristic information needed for particular purposes:

- o *Plant Development.* Specific information on plant growth and development features would be extremely useful. Recommendations for additional data include:
  - Growth characteristics (e.g., size, "weediness")
  - Estimation of horticultural potential
  - Environmental amplitude beyond that in natural habitat
  - Energy content (e.g., wood, resins)
  - Location of ecotypes for genetic development

- o *Management/Control.*
  - Host information for pests
  - Known diseases
  - Degree of weediness
  - Rate of spread of weeds
  - Life cycle of pests/weeds (including dispersal mechanisms)
  - How weeds behave in various parts of range (e.g., Florida vs. California) and where they occur in crops (e.g., in row crops, turf)
  - Information on escaped species that are not yet weeds
  - Potential weediness information
  - Conservation status

- o *Plant utilization.*
  - Forage value of crop species
  - Edibility/toxicity
  - References to uses of plant materials

IV. REQUIREMENTS FOR MAKING FLORISTIC INFORMATION MORE USEFUL

Our group identified a number of factors that should be considered in order to increase the overall usefulness of FNA, in addition to the additional data fields suggested above.

A. *Ease of Use*. FNA must be easy to use, particularly with respect to identifying taxa. Features that would aid the non-taxonomist in using FNA include:

- Vegetative keys and descriptions
- Keys that "work" in all stages of plant development
- "Key" diagnostic features emphasized in descriptions
- Descriptive information increased
- Metric/English measurements used
- Nomenclatural changes (synonymy) clearly explained
- Precise distribution information (e.g., latitude/longitude) provided
- A user manual/pronunciation guide provided
- Color photos, illustrations, habitat photos, and/or other sources of illustrations provided

B. *Consistency, reliability, and standardization*.

- Data entries in FNA must be as complete as possible and consistent
- Each user group must be confident of the accuracy and reliability of the information in FNA
- Standardization of geographic names is particularly important

C. *Electronic database*. It is clear that FNA most likely will exist in an electronic form as well as a series of printed books. Any electronic database of information also must be easy to use, widely available, and available at low cost to the end user. The FNA computer-based database should be cross-linked to as many existing databases as is feasible.

D. *Nomenclatural/Bibliographic Considerations*. It is particularly important to the non-taxonomist user of FNA that it use standardized scientific and common names as much as possible. Full citations of authorities also would be helpful. In addition, FNA must be fully documented to primary literature sources (including an extensive bibliography) and some indication should be included as to how current the information in FNA is.

E. *Geographical/Habitat Data*. The following kinds of geographical/habitat data should be included in FNA:

- Detailed information on soils (e.g., parent material)
- Information concerning associated plant communities
- Natural range of taxa as well as present range
- Detailed habitat descriptions
- Suggestions as to occurrence outside known range
- Distinction between native occurrence and naturalized distribution

V. NEW APPLICATIONS OF FLORISTIC INFORMATION.

There is no doubt that a Flora of North America that included the kinds of information outlined above would stimulate and support new research efforts in agriculture, forestry, and horticulture. In addition, FNA will be an important resource tool in teaching and extension programs throughout the United States, as well as in commercial applications. Our group identified a number of potential areas where FNA would be extremely useful to agricultural/horticultural research, teaching, and extension programs. These can be grouped into the following general categories:

- Floristics (local Floras/guides)
- Resource management/control
- Conservation efforts
- Education (especially in extension activities)
- Development of new products/crops/ornamentals
- Phylogenetic/evolutionary studies

VI. RECOMMENDATIONS AND CONCLUSIONS

In order for FNA to be used broadly in agriculture, forestry, and horticulture, there are several minimum requirements for the project:

- FNA must be a "distribution-based" flora; i.e., precise distribution data must be included for each taxon included in the flora.

- FNA **must** include phylogenetic information. There must be some type of information about the phylogenetic relationships of taxa, even if it is a "best guess" of an expert. These data are of particular importance to individuals working in the crop development area.

- FNA must be easy to use. It needs to be as non-technical as possible, as long as this simplicity does not lead to ambiguities. There was consensus among our group that this will lead to greater usefulness of FNA.

- FNA must be well-illustrated and/or provide documented sources of illustrations for each taxon discussed.

- FNA **must** be well-documented, including an extensive bibliography to primary data sources. Users must have confidence in the accuracy of the information included in FNA.

- Data included in FNA must be presented in a consistent fashion.

Our group clearly recognized the need for FNA to exist in a printed form as well as in an electronic form (i.e., computer databases). The printed form is essential to maximize the use of FNA. Our group recommends that several new databases be developed to include the kinds of information outlined throughout this report.

However, it is important that FNA not duplicate existing databases. Instead, it is critical to the success of FNA that any new databases developed be cross-linked to existing databases that are applicable. Such databases must be networked and linked to other similar databases on a national and international level. It is also critical that the electronic form of FNA be widely available at low cost to the end user.

APPENDIX 2

Report of Group 2. How is floristic information used to assess environmental conditions? How can this be improved, or increased?

Group Members

Barkley, T. M. (recorder)*
Brigham, W. U.
Graham, R. L. *
Hayne, D. W.
Macrander, A. M.
Magill, R. E.
Morain, S. A.*
Simpson, B. B. (leader)*
Stein, B.
Waggoner, G.

I. INTRODUCTION

Implied throughout the discussion was the fact that the FNA has two parts: the production of a Flora in a series of volumes; and the simultaneous production of a multifile botanical database with ongoing, expandable capabilities, particularly one having Geographical Information System (GIS) capabilities. The discussions largely centered around the building and utilization of this database. The group believes this database would be highly applicable to the mission of federal, state, and local government agencies and private organizations having land resource and environmental responsibilities (e.g., BLM, Nature Conservancy, timber companies, environmental consultants).

II. USE OF FLORISTIC INFORMATION IN ENVIRONMENTAL ASSESSMENT

Utilization of current Floras among participants has varied from "considerable" to "little." There was universal agreement that Floras are used for identifying plants and determining their correct names; and, in a more general sense, as guides to habitats and plant associations. Detracting from the utility of available, published Floras are the often clumsy keys, cryptic nomenclatural citations, and far too general information on habitats and distribution.

III. REQUIREMENTS FOR MAKING FLORISTIC INFORMATION MORE USEFUL

All agreed that the printed FNA and its companion database will be more useful to a broad community if they contain more information than do currently available Floras, and if these data are more easily accessible. The following were regarded as minimum requirements:

- Illustrations (especially line drawings)
- Data on relative abundance or rarity

---

* Asterisks indicate members of the writing team for the group

- o Status as threatened or endangered
- o Economic significance
- o Plant/animal associations
- o Ethnobotanical uses

Other data desired by group 2 are given in Table 8. The precision of the data, particularly locality information, should be increased. Eventually the database should include a specimen-based file with good locality data. It was generally presumed that the utility of the databases will increase as the amount of information contained in them increases.

IV. NEW APPLICATIONS

If this expanded floristic information were easily available, it would make it possible to answer questions that could not have been answered in the past. First, the database could be used, for instance, to determine the locations of threatened and endangered species, habitats of notable botanical diversity, chromosome races, or changes of distributions over time.

Second, it would permit the correlation of plant distributions to other spatial data such as soil types, climatic parameters, cytological races, historical land use, topography, and, potentially, animal distributions. Such correlations would permit defensible studies of landscape evolution and ecology.

Third, such information would enhance our ability to make environmental predictions and sound siting decisions.

There was a definite opinion that the designers of the database should make the system "user friendly" (e.g., few commands, menus, an understandable manual). It was an article of faith among the discussants that the technology of databases is advancing with sufficient speed to accommodate the needs of potential users and to present floristic information in any form desired (text, flat files, polyclaves).

V. RECOMMENDATIONS AND CONCLUSIONS

The overall discussions of the group centered primarily around the database. The published Flora, it was agreed, should be user friendly and convenient for identification and nomenclatural matters. For use in environmental assessment projects, a specimen based database will be the most important component of FNA. The database will be a primary element in addressing spatial issues, networking with other databases, and generating significant layers for the development of Geographical Information Systems.

The need for reporting known non-occurrences of species across space might be considered in the future. How this information might be collected and entered into a database needs further investigation. The need for such information should diminish as the accuracy of positive occurrence data increases.

TABLE 8. INFORMATION TO BE INCLUDED IN FNA

ENVIRONMENTAL TOPICS

- Precise locality data
- Plant communities and associations
- Soil types
- Uses--commercial, ethnic
- Seral stage
- Ancient abundance over range
- Management (current and historical)
  --beneficial or deleterious
- Ethnic perspective
- Phenology
- Weediness--response to perturbations
- Pollination/seed dispersal
- Herbivores and distributors
- Spatial and temporal distribution
- Phytochemistry
- Spectral properties of dominant taxa

TAXONOMY

- Standardized nomenclature
- Native versus exotic
- Accurate descriptions that agree with keys
- Brief diagnostic statement
- Chromosome number, races
- Regional common names
- Nomenclature--types
- Annotated bibliography
- Hybridization

EXCURSION MANUAL NEEDS

- User-friendly keys
- Detailed distribution maps

APPENDIX 3

Report of Group 3. Use of floristic information to assess organisms' responses to changes in environmental conditions, e.g., expansion or contraction of ranges, vitality of populations, genetic changes.

Group Members

Argus, G. W.
Brouillet, L. (recorder)*
Fay, J. (leader)*
Flint, H.*
Gibson, A.
Heywood, V. H.
MacBryde, B.
Morse, L.
Murray, D.
Shetler, S.
Walter, K.
Whetstone, R. D.

I. INTRODUCTION

Before attempting to address the topic, the group discussed the nature of floristic information. Two interpretations were advanced:

- The first one restricted the definition to that information that is traditionally tied to a Flora, i.e., taxonomic (nomenclature, classification, literature), morphological (including identification aids), and distributional (geography and habitat).

- The second one considered that floristic information includes any datum that is tied to taxa belonging to a specified area.

No consensus was reached. Nevertheless, the group found it possible to address the question in historical perspective.

II. USE OF FLORISTIC INFORMATION TO ASSESS ADAPTATIONS

The major conclusion reached was that conventional floristic information provides a very limited basis on which to assess the responses of organisms to changes in environmental conditions.

Conventional floristic information provides a background. Although individual Floras tend to be static, a set of Floras may sometimes give indications of dynamic processes over space or time. If we include herbarium data within the concept of floristic information, it might be possible, for instance, to document the continued existence of a given population over periods of time, or the range expansion or contraction of a taxon.

---

* Asterisks indicate members of the writing team for the group

III. REQUIREMENTS FOR MAKING FLORISTIC INFORMATION MORE USEFUL

To answer questions about change, we need to know what is present at a given place at a particular time. If we want to make floristic information useful in monitoring change, the group recommends:

1. That the classes of data gathered and presented be broadened and recorded in a systematic manner. To do so, *intensive field-work* is needed at the *population and site levels*, so that the types of data listed in Table 9 can be gathered, on both a *spatial* and *temporal* scale, with emphasis on *variability* within and between sites and populations. Special attention should be given to marginal populations.

As a result of this finer-grained data set, we will be better able to provide the following information:

- o More precise phytogeographical descriptions, particularly of edges of ranges, and elevational limits
- o Better knowledge of geographic history of taxa (expansion, contraction, introduction, etc.), and better capacity to document future changes
- o Phenological data across ranges
- o Better description of variability within and between populations, and of changes over time in these parameters
- o More precise data on ecological requirements of species over their range

At a technical level, it might become necessary:

- o To develop better means of recording phenology across a range
- o To use or develop instruments and tests to measure physiological and chemical characters in the field
- o To take on-site photographs (which could be digitized in databanks)
- o To consult with ecophysiologists before studying micro-climatic parameters, so that the data collected might be significant to them

2. That FNA adopt an organized grid system to describe geographical data.

TABLE 9. ADDITIONAL FLORISTIC INFORMATION TO BE GATHERED FOR INCLUSION IN FNA

BIOTIC FACTORS

- Innate attributes
- Breeding system
- Expanded morphological description
- Life-form, underground parts, juvenile stages, pera- and paedomorphosis
- Other labile traits, e.g., odors, etc.
- Phytochemistry
- Diaspores and dispersability
- Germination and establishment
- Resilience to stress
- Invasiveness

ABIOTIC FACTORS

- Soil/substrate
- Geological history
- Elevation
- Disturbances (physical)
- Microclimates

POPULATION/COMMUNITY ATTRIBUTES

- Demographics, trends
- Genetic diversity
- Phenological data
- Hybridization
- Frequency of morphs and clinal variation, etc.
- Biotic interactions: pollination, parasitism, synecology, and other plant-animal (etc.) relationships (site, species)
- Toxicity
- Weediness

ANTHROPOLOGICAL FACTORS

- Values/Uses of taxa (ethnobotany, economic attributes, etc.)
- Land-use history
- Overcollection (botanical and economic)
- Conservation status
- Evidence of stress
- Seed storage
- History of introduction

3. That databases and networks be used to store the broadened data set to increase the usefulness of floristic information in answering questions about changes.

- o Particular attention should be devoted to Geographic Information Systems.
- o Links should be established with existing databases of the following subjects:
  - Geology and geological history
  - Climate
  - Vegetation
  - Nomenclature, synonymy, etc.

4. The style, content, and format of the information be tailored to the needs of the various users. In order to insure that this is done, a survey of the needs of user groups should be conducted. The following characteristics should be incorporated into the design of the books or database:

- o Information should be presented in a consistent manner, comparable across taxa
- o Data should be made readily and widely available to users, for example, via computer identification systems
- o Flexibility must be built into the database because technology is changing and new fields might be needed

IV. NEW APPLICATIONS

Overview

Continental floristic information provides a powerful resource for comparative questions and applications at different scales among units. Questions and applications can be made within modules (components of the floristic data), among these modules, and across multiple data centers each with their modules (combining flora data with other data). Also, noncomparative questions and applications can draw directly upon floristic data. In addition, more specialized and utilitarian derivative data could be produced from, or in coordination with, the Flora project.

A continental Flora provides the means to approach thinking globally, while acting locally. Research questions can be better formulated and applications accomplished when placed in broad context, making local awareness and activity more significant and effective. Degrees of confidence in information can be assessed, gaps determined, and "remedial" research conducted. Such a tool becomes the *de facto* standard or baseline for current use and future effort (as well as reflection to the past), while at the same time continually challenging and stimulating new work. To maintain its effectiveness over time, this *de facto* standard

around which other work is organized requires frequent periodic revision coordinated among the major users of the taxonomic data standard.

Used alone, the continental floristic information can answer a multitude of questions concerning the plants that occur in the continental region, in this instance the vascular flora of North America north of Mexico. Drawing upon other major floristic information resources (e.g., Europe, China), the North American data can help answer questions of inter- and multi-continental scope. Combined with other taxon-oriented databases, deeper research questions can be explored. Combined with other geographical data systems, patterns can be sought and biogeographical questions approached. Combined with specimen-label data networks, more detailed geographical, ecological, and historical aspects can be considered. Combined with databases on conservation status, land-management responsibilities, etc., the floristic information can be applied to natural resource problems. Combined with other concerns of society, such as tourism/recreation, pride in locale, economic potentials, etc., floristic information can be made more relevant and the importance and value of plants recognized and enhanced.

Structure of floristic data

The major kinds of floristic and related information, in the broad sense, include:

- Taxonomy (identification, description, classification, and nomenclature)
- Geography and habitat
- Life history/species biology
- Interactions with abiotic factors and with other organisms [nonhuman (natural) and human]

Added to this are bibliographies and other sources of information. This information, and the concerns it addresses, relate to: what is it, where is it, what does it do through time, and what (natural or human) interacts with it through time? Data in depth, well organized into these modules, becomes so quantitatively thorough that qualitatively different perspectives can emerge.

Taxonomic, descriptive, and identification data are essential components of a floristic information synthesis used in ecological and conservation studies. Indeed, it is the taxonomic classification that forms not only the framework or organizational structure for the published Flora and the floristic databases, but also for numerous coordinated or independent databases containing species-related information that can be consulted in conjunction with the floristic data *per se*. Providing this taxonomic classification in an information-management framework is a fundamental role of systematic biology. To the extent that the systematic community does not provide its information properly synthesized and accessible, other information and activity will develop less soundly than it should, with loss for all concerned (both people and plants).

This botanical classification also forms the major linkage between information about the plants as plants, and information about these plants and their interactions with other kinds of plants (e.g., vegetation ecology) or animals (e.g., pollinators or herbivores). The availability of a continental classification also reduces provincialism and encourages greater consistency among the classifications employed in local and regional Floras, field guides, vegetation or wildlife studies,

environmental assessments, and other kinds of natural-history research that include the names of plants.

The taxonomic organization of the continental floristic information also provides the capability for determining pairs of related taxa that can be studied in comparative studies, for example in morphology, physiology, or anatomy. Also, in horticulture and conservation biology, one may need to develop and practice techniques on an abundant, related surrogate before working directly with an endangered entity.

Furthermore, the production and publication of a Flora of continental scope will provide an assessment of the state of the taxonomic and floristic knowledge of the species, bringing the information on each up to a scientific standard, thereby giving conservationists and resource managers greater confidence that local endemics are distinct entities rather than misidentified odd variants or wide disjuncts of other taxa. Noting infraspecific variation (taxonomically recognized or not) in the floristic information will also increase awareness of the pattern of variation in a species, helping in determining which of its occurrences are most significant for conservation. If the information on variation can be combined with ecological or geographical data, many other kinds of studies would be facilitated.

Identification of individual plants is essential to most kinds of botanical research, including agriculture, horticulture, forestry, and plant conservation. While local and regional Floras and other specialized guides will continue to be used in the vast majority of routine cases, the availability of an identification manual on a continental scale is important when a plant is found that cannot be readily identified by works of more limited scope. Access to this continental perspective is particularly important in distinguishing wide disjuncts from overlooked endemics or from undetected escaped exotics. Availability of the information as polyclaves (multiple-entry keys) would greatly facilitate its utility, especially for identification of incomplete or vegetative specimens, such as those often encountered in ecological studies.

Species biology and other natural-history characterizations, including phenology, form a major unit of information important to studies of distributions, habitat relations, and conservation needs. Phenology, for example, is important information in planning field research, either in deciding when to seek a particular species, or what species may be appropriate to seek in an area at a particular time. However, little information is conveyed by generalized continental-level phenological data for widespread species; state or finer degrees of precision are essential. Combined with geography, ecology, or other factors, this phenological data could also be used in research questions relating phenological patterns to such other factors.

Similarly, information on breeding systems or other parameters of species biology, if collected and presented in a standard form for the continental Flora, could be used in scientific studies, for example to seek similar patterns in disparate taxonomic groups, different habitats, or different geographic regions. One interesting possibility is to seek particular patterns of life-history parameters that characterize rare or declining species. Management techniques developed for a particular species, whether a rare native or a problem exotic, might also be of interest for management of biologically similar species, regardless of their taxonomic placement or geographic location. Many of these characterization

parameters would also be of interest to nature interpretation or other activities involving the general public.

Biogeographical data is one of the most important categories in a Flora, and, in a more general sense, geographical data is one of the most powerful kinds of information with which floristic knowledge can be compared. Traditionally, floristic information beyond the local level is in summarized form, usually by political jurisdictions, natural regions, and similar polygons. Grid-cell data is sometimes used, but requires substantial effort to produce, even on a coarse scale, since the source data often do not appear that way. Detailed specimen-label data provide even more powerful geographical perspectives, when used with the proper cautions regarding misidentification, collecting histories, and poorly known areas.

Geographical analysis of the floristic data, even a simple itemization of states for species, can focus research by showing gaps in reported distributions and areas in which further fieldwork would seem important. Such techniques are far more powerful when done to the county or similar level of resolution. This kind of focused field effort is particularly important in determining current distributions for rare species and for assessing present and anticipated future distributions of problem exotics. The geographical data also provide a powerful tool in initial screening of species that may be rare, declining, or endangered, although current fieldwork is then necessary to determine their actual status. Similarly, these data can be used to identify centers of endemism or disjunction, which, if not already well studied, may be appropriate priorities for further field research and collecting as well as for conservation review. These data would also assist systematists in locating field-study areas in which particular combinations of species might be found growing sympatrically or allopatrically.

Combined with the many kinds of geographical, geological, climatological, and other data now becoming available in geographical information systems (GISs), the floristic data, when of sufficient resolution, can be analyzed in numerous powerful ways, such as examining factors that may determine distribution limits, or identifying specific places in which particular species might be sought. GIS-based research might also address biogeographic questions such as whether particular distribution patterns are environmentally or historically based (including both geological history and human history). For an aggressive exotic, analysis using a GIS might indicate the extent of the potential range and the areas in which it might become a management problem.

However, it must be noted that the degree of precision needed for effective GIS work is rarely available, except for well-studied rarities. Even distribution data at a precision of, say, a quarter-degree square would require tremendous amounts of new field research for the vast majority of North American vascular plants. Nor can GIS methodology as presently available be used to assess long-term historical patterns in range changes, since the available specimen label information and other historical data are insufficient to indicate places in which plants formerly did not occur.

Specimen labels and related historical accounts (such as old local Floras) are the major information sources for specific instances of changes in range, particularly for sites in which a species formerly occurred but cannot be found now. (The problem of showing a species did not occur at a given place in the past is much more difficult, but can sometimes be made by concluding that the area was sufficiently well known, as shown by other collections, that the plant in question

would have been collected had it been there then.) Availability of specimen label data in an organized form would enhance botanical research in many obvious ways, generally by increasing the awareness among taxonomic and floristic researchers of the pertinent specimen holdings of the many herbaria. Ecologists, conservation biologists, and others would also regularly consult such a resource if it were available. The specimen data could also be used to investigate the depth and adequacy of the collecting record for a species, or to analyze the relative degree of collecting activity in various areas at various times.

Habitat information, and information on interactions with other organisms and with abiotic factors such as substrates or microclimate, form a fourth major category of continental floristic information. Included here are the interactions, both positive and negative, with people and with human activities. Potentially, this is one of the most powerful capabilities of the published Flora and database, but traditionally it has been one of the weakest components of floristic work. It is in this area that those outside plant taxonomy seek botanical information for *their* purposes, and where the writers and editors of Floras and producers of floristic databases can communicate the wealth of botanical information deemed important. The database, even more than the published Flora, becomes more important here, and this area can be expected to grow considerably in the subsequent more enriched floristic database system.

One major use of habitat data is in assisting field researchers to determine the particular kinds of places in which a species might be found. For conservationists, habitat data help determine the kinds of sites offering the best opportunities for long-term protection and management of a species. Relating habitats to life forms or other species-biology factors should also be a productive avenue of data analysis and may suggest many new research possibilities. Finally, characterization of habitats of rare species or problem exotics should help land managers and environmental planners determine which species might be affected (positively or negatively) by particular vegetation-management practices or land-use changes.

V. REQUIREMENTS FOR MAKING FLORISTIC INFORMATION MORE USEFUL

Several ways were suggested to make the data useful. These can be divided into general and special approaches:

A. *General approaches*. Below we follow a roughly chronological sequence in the development of the product, i.e. a Flora or a database.

- o A first step should include the identification of potential audiences and the testing of prototypes for product utility among these particular users.

- o In designing the "system," the following considerations should be addressed:

  - access to databases should be made as inexpensive as possible
  - complexity of presentation should be tailored to the intended audience
  - interfacing should be made consistent throughout the system
  - information, including that stored in electronic form, should be citable

- keys should be illustrated as appropriate
- the use of colors should be considered whenever useful

o The following classes of end-products and services should be envisaged:

- hard copies tailored to particular user groups
- extension and answer service
- periodic publication of indices to the content of databases
- training of consumers to use the systems

B. *Special approaches*. Among these approaches were standardization of common names and production of specialized versions of the Flora for handicapped groups.

C. *Other approaches*. These were essentially technological, for instance:

o expert systems approaches to particular needs
o chemical identification kits
o multiple-entry computer keys
o cassette storage of information

A general conclusion to this section is that the system should be sensitive to the needs of its users.

V. RECOMMENDATIONS AND CONCLUSIONS

Existing floristic information can be used only in a limited way to assess organismal responses to environmental changes. New classes of data must be gathered and addressed at a finer scale, particularly at the populational and site levels, if floristic information is to become more useful in assessing dynamic processes. More field-work, as well as specimen-data capture, are needed to accomplish this. Databases and networks will render this type of analysis possible. Consumer needs should determine the design of information storage and presentation.

APPENDIX 4

Report of Group 4. Uses of Floristic Information in the Biological Sciences other than Plant Systematics

Group Members

Beecher, C.
Catling, P. M.*
Gibbs Russell, G. E.*
Hickman, J. C.*
Johnston, M. C.
Kartesz, J.
Kosztarab, M.
Richards, J. H.
Shultz, L. M (recorder)*
Stuessy, T. F. (leader)*
Taylor, I. E. P.*

I. INTRODUCTION

One of the most important considerations in development of the FNA project is to ensure that the data contained in published volumes and in any associated databases are maximally useful. This section of the workshop report is concerned with the kinds of information that should be contained in the FNA project. We outline below the nature of floristic data, the subdisciplines of biology which make up the non-taxonomic user community of FNA, the ways in which these subdisciplines currently use floristic data, new types of data or ways of utilizing these data that these subdisciplines would like to find in FNA, ways to gather and present these new data, and new research questions that might be asked of this expanded data base.

II. USE OF FLORISTIC INFORMATION IN THE BIOLOGICAL SCIENCES

A. The nature of floristic and monographic data:

To discuss what types of floristic data are used in the biological sciences outside of plant taxonomy requires an understanding of what floristic data are. At an elementary level, floristic data deal simply with what plants grow in particular geographic regions. At the other extreme, floristic data encompass all data about plants growing in particular areas. For purposes of FNA, we believe it useful to think in terms of "minimal floristic data" and "monographic data." *Minimal floristic data* consists of:

- o the name of the plant (scientific and common)
- o its geographic location (where it grows)
- o its morphology (what it looks like)

---

* Asterisks indicate members of the writing team for the group

- how it differs from other plants (especially as revealed in keys and diagnostic illustrations)
- flowering time
- habitat
- elevation

Other data (Table 10) about plants are regarded here as revisionary (or monographic) data, some of which are included in floristic treatments (e.g., chromosome numbers), but most of which are not. One of the challenges of FNA is to decide which of the revisionary data might be included in the published volumes or databases.

B. Subdisciplines of biological sciences and their dependence upon floristic data:

Biology is an extremely broad and complex scientific discipline that encompasses the study of all life and its processes. To provide a focus on how the different subdisciplines of biology use floristic information at the present time and how they might use other (revisionary) information emanating from the FNA project, a brief discussion of these areas (Table 4, p. 62) is useful.

Many of the areas of biology focus primarily on the types of data such as anatomy, biochemistry, or cytology. In these subdisciplines, which can be called data-oriented areas, questions asked are those that can be answered by single types of data rather than depending upon data of many different types. Other areas of biology, however, depend routinely on several or many different types of data, which can be called question-oriented areas, such as evolutionary biology, molecular biology, and systematic biology.

The large body of floristic information that currently exists is, to a large extent, not accessible to non-taxonomists. There are many reasons, the most notable being:

- inconsistent treatments
- a general lack of dynamism in the teaching of floristics
- general difficulties arising from confused nomenclature and lack of knowledge about how to obtain information

It is likely that the Flora of North America will unify nomenclature and treatments. The challenge for taxonomy is to draw the users in and to encourage a commitment to follow through until information is obtained.

Users of floristic and revisionary data to be contained within the FNA published volumes and databases exist in both of these major classes of biological disciplines. Some areas now routinely use minimal floristic data, such as ecology and population biology, and others do not (e.g., biophysics). All subdisciplines, however, would definitely use some minimal floristic data and much more of the revisionary data if they were readily available. Some of the existing demands presently being made for floristic data by non-taxonomic users are shown in Table 3 (p. 61).

TABLE 10. TYPES OF REVISIONARY (OR MONOGRAPHIC) DATA USED IN PLANT TAXONOMIC RESEARCH

FROM PLANTS THEMSELVES

- Structures
  - Anatomy
  - Chemistry
  - Cytogenetics--cytology, chromosomes
  - Embryology
  - Genetics
  - Morphology
  - Palynology
- Processes
  - Physiology
  - Biochemistry

FROM PLANT-ENVIRONMENT INTERACTIONS

- Abiotic
  - Soils
  - Climate
- Biotic
  - Predator-prey
  - Pollination
  - Dispersal
  - Symbiosis

C. Other Users of Floristic Data

Many people outside of the biological sciences also use floristic data. These people may have a biological background (e.g., ecological consultant) or they may be dealing with floristic data as their first exposure to biology (e.g., mining engineer).

Table 5 (page 65) presents questions and concerns that might be posed by a selected sample of these users. It is important and perhaps surprising to note that the answers to almost all of these questions can be obtained, at least in part, from current forms of Floras. However, it is likely that most potential users do not know either that Floras contain the required information or how to abstract it from a Flora.

III. REQUIREMENTS FOR IMPROVING THE USE OF FLORISTIC INFORMATION

A. New types of floristic data and their uses:

The functioning of the FNA project presents an immense opportunity to improve the consistency of many kinds of data on the plants of North America and to increase access to those data substantially. Although the published volumes and the databases will be constrained in different ways, it is desirable to consider the ideal conditions for diverse users outside the fields of floristics and plant taxonomy. The published volumes will be considered separately from the databases.

Space and format constrain the published volumes. Despite the inflexibility of the format, these volumes will provide the only possible access to FNA information for many potential users (at least in the near future). Therefore, as much information as possible should be published. Inclusive (but not exhaustive) lists of information types are given in Table 11.

We assume that most kinds of data listed in Table 11 can neither be published nor be included in the primary database here assumed to correspond closely with the published volumes. Data that cannot be so accommodated should be entered into the extended database as soon as they can be compiled. It is from the extended database(s) that most non-taxonomists will derive FNA information that is important to them and not currently available.

Because of the diversity of users, much attention must be paid to the ease with which all FNA resources can be used. For both the published volumes and the databases, it is crucial to create "user-friendly" formats and environments, especially to keep plant names current but cross-referenced with synonyms under which information might be sought.

To improve the usability of the published volume, we offer the following suggestions:

- standardize use of common names
- construct workable keys
- provide maps of distribution ranges

- o differentiate misapplied names from nomenclatural and taxonomic synonyms
- o publish a field-portable, single-volume identification manual (perhaps keys only)

To improve usability of the database, we suggest the following most important ideas:

- o standardize database structure and presentation
- o construct fully relational, encyclopedic databases with capacity for additional fields
- o provide information about and access to other databases
- o provide phone-linkage access
- o provide on-line identification via multiple-entry keys
- o construct a bibliographic database

Further, it would be helpful to:

- o point to gaps in data
- o distribute disks
- o develop front-end artificial intelligence capability
- o provide more graphics, including illustrations
- o stimulate computer ownership and literacy
- o provide a hierarchical, numerical code for each taxon

B. The gathering and presenting of new floristic information:

Much information about the plants of North America already exists in the literature (e.g., plant physiology, biochemistry, biophysics). This information needs installation into its own database, which in turn should be compatible with the FNA database. This type of information is thus not new, but production of a database would make it new to the floristic world and available for correlative research.

It is likely that these new databases will be very incomplete. The fact that this incompleteness can be detected will indicate that the framework is ready to receive truly new information. Once databases are developed the first questions can be asked and answered. These first questions will be bibliographic and correlative. Beyond this, future uses are not easy to predict and must be continuously considered as FNA develops.

As the FNA project progresses, the databases set up in parallel with it will become a developing resource, independent of the publications derived from it. When publication of each FNA volume has been achieved, the database underlying it will continue to grow: through updating the published data; through expanding the depth of information in the basic floristic data; and through adding additional kinds of information not included in the published volumes.

TABLE 11. ADDITIONAL DATA TO BE INCLUDED IN THE PUBLISHED VOLUMES OF FNA

INFORMATION ABOUT PLANTS THEMSELVES

Structural Information
- Chemicals produced
- Toxicity
- Chromosomes
- Seeds
- Embryos
- Seedlings
- Pollen
- Buds
- Wood
- Overall variability among plants
- Variation in structure with environment, age, position, etc. (phenotypic plasticity, heteromorphy, etc.)

Processes
- Patterns of polyploidy
- Seed dormancy
- Developmental patterns and sequences
- Detailed phenology
- Breeding systems

INFORMATION ON ENVIRONMENTS AND INTERACTIONS

Environmental Properties
- Geological substrates
- Soil (especially specificity of indicator status)
- Distributional details (resolution better than state level)
- Elevation ranges
- Climatic data

Biotic Interactions
- Important pollinators
- Dispersal vectors
- Herbivores (especially insects)
- Parasites
- Pathogens
- Forage values
- Crop (or other horticultural) potential

Information in the database will be available in several ways:

- Subsets of the whole Flora of North America, such as an abbreviated manual for identification only, or computerized information for interactive identifications, covering the whole area; or the complete information for particular areas, such as regional, state, or local floras

- Correlations of particular fields--any data element in conjunction with any other, or presented in such a way that FNA data can be correlated with data from other databases. This use may be expected to grow as information becomes more complete, and especially as additional information of biological significance is added to the basic floristic data.

Information of all types should be made available in hardcopy of various kinds, and electronically, both as a basis for further publications and as input for related computerized databases. Electronic communications will change as technology develops, but now include diskette, tape, compact disk, and dial-up access. Particular attention should be paid to international data transfer formats set up by the Taxonomic Databases Working Group (TDWG), and FNA should be involved in their development.

Published volumes continuing and ancillary to the Flora of North America that are derived from the computer system will be limited only by the imagination of FNA users. Possibilities include future editions in the present or expanded format, regional and local floras and manuals, as well as greatly expanded treatments of specialized aspects through correlations of the basic floristic data augmented by more biological information in specialized groups.

IV. NEW APPLICATIONS

The specific applications and research-oriented questions resulting from a current continental Flora are vast in number. In general, Flora of North America will provide a wealth of current information on North American plants to biologists, students, government agencies, planners, and indirectly to the public at large. In the process of producing the systematic treatments for the flora, areas of taxonomic complexity will be highlighted and hypotheses will be generated. Each completed volume will provide answers to a multitude of questions, but each volume will also stimulate many new questions. The applications and inspired research may be conveniently considered under three headings, with several examples provided for each.

A. Descriptive information:

A continental Flora will provide an accepted system of naming plants that is essential to communication in scientific research, education, and management.

The volumes of the Flora will provide an essential means of plant identification, but there is potential for the database to provide identifications also through polyclaves and digitized images. It is anticipated that it will become possible to have color slides identified by the computer.

Specific morphological and other information will be available to answer questions such as: What are the seeds of a certain plant like? When does it flower? What is its habitat? Is it poisonous? Diagnostic morphological structures, such as species-specific insect galls, will be described. The associated database may be able to answer questions regarding chemical and physiological characteristics, tissue culture, genetic basis of variation, population structure, hybridization data, transplant experiments, chloroplast DNA, enzyme patterns, phylogenetic constraints on evolutionary change, etc., sometimes through networking with other databases. It will be possible for a biophysicist to obtain from FNA information on growth rate of various tree species in order to obtain material with different growth rates to determine the relationship between growth rate and wood characteristics. Descriptive morphological (and other) information will allow an analysis of co-evolutionary patterns.

Because of its current and authoritative nature, the *Flora of North America* (including the computerized database) will serve as a starting point for the study of systematic problems involving all North American plants. Many of these problems will be elucidated during the production of the Flora. FNA will also encourage the development of systematic-oriented hypotheses, and the testing of these hypotheses to the extent that the required information is available.

Using the *Flora of North America*, it will be possible to know the status of various species, whether they are rare or endangered, whether they are native or introduced, whether they are noxious weeds, etc. Working with the data base, it will be possible to discover and monitor changes in status, such as the decline of native flora, the spread of alien weedy species, and perhaps even changes in the size of populations. Hypotheses may be generated to explain the observed trends.

Through comparison with other temperate region Floras, FNA will contribute to a better understanding of the flora of north temperate regions around the world.

From a continental Flora and associated database, it will be possible to extract floristic treatments of specific regions, including states, provinces, territories, parks, or vegetation zones. It will also be possible to determine vegetation zones on the basis of species distributions. Specialized lists and other information may also be extracted (sometimes through networking with other databases), such as regional lists of poisonous plants, lists of plants containing certain chemical compounds, etc.

B. Predictive information:

Predictions can result from any of the categories (e.g., phylogeny, distribution, flowering time, historical aspects) of information in the volumes of the flora or the database. For example:

Certain secondary metabolites, chemical compounds, or kinds of toxicity may be expected in particular species on the basis of occurrences in related species.

If a plant occurs in a specific habitat or within a specific region, its occurrence at a given site in or near that region may be predicted.

Information on flowering time permits a variety of predictions having to do with aspects such as community dynamics and management. A major discomfort to hay fever sufferers will coincide with flowering of ragweed in a particular region. The return of broad-tailed hummingbirds to parts of the Rocky Mountains in spring may be predicted to coincide with the first flowering of plants with reddish, tubular flowers.

Patterns of invasion of newly introduced and other categories of plants may be anticipated on the basis of past invasions and relationships in terms of morphology, ecology, strategy, etc.

The distribution of phytophagous insects may be predicted on the basis of host plant distributions; forecasting the distribution of pest insects will be facilitated through additional use of climatic data.

C. Correlative information:

The possibilities here extend to the limits of the imagination. Correlative information demonstrates trends and results in new hypotheses. It will be generated largely from the database and through networking with other databases, rather than directly from hard copy of the Flora. There may be very practical reasons for seeking correlations, or they may be academic and hypothesis-generating with practical spin-offs. For example:

What is the correlation between the distribution of species indicative of high-quality cattle range and present cattle density, to suggest regions of potential rangeland?

What correlation exists between the geographic occurrence of rare species and protected areas, to determine whether there are substantial concentrations of rare species in need of protection and to assist in prioritization of areas for protection? It will also be possible to correlate known distributions of rare species with environmental data, to predict new localities or to find appropriate sites for re-introduction.

What correlation is there between species diversity and vegetation zones or urban areas, etc.? A practical spin-off may be highlighting of the region or zone where maximum biological diversity can be preserved, or of a region that is a logical one for urban development as a consequence of low biological diversity.

Other examples include the correlation of: aspects of breeding system (inbreeding versus outbreeding, heterostyly with homostyly, etc.) with geographic distribution, elevation, overall environment, etc.; flowering time with geography; geographic distribution with pH and edaphic factors; developmental or physiological features with the environment; cytology and geography; extent of range with environmental factors; higher categories with environmental factors; alpha and beta diversity with vegetation type; variability in taxa with regard to vegetation zones, environment, etc.; and broad life-form categories with geographic distribution and environment.

V. RECOMMENDATIONS AND CONCLUSIONS

- Minimal floristic data to be included in FNA include name, location, morphology, key characters, flowering time, and habitat. Other data should be included insofar as possible. Minimal data and some kinds of additional data reflect the general demands of floristic data made most often by both systematists and non-systematists.

- Different researchers in different subdisciplines of biology use floristic data to varying degrees but some depend heavily on it. All would use

some minimal floristic data and there would probably be even greater use of revisionary data if they were readily available.

- At least 45 different career specialties outside of plant systematics utilize floristic information routinely. The scope of activities into which plant biologists (and especially those needing floristic information) are being drawn is ever-expanding. In addition numerous subdivisions of applied biology and other disciplines also depend on floristic data.

- No matter how inexperienced the user, he or she must be drawn into the system and led to the core data in a way that compels him/her to present questions in a way that will deliver a useful answer.

- The use of FNA may be increased if information on toxicity, chromosomes, seed characteristics, overall variability, patterns of polyploidy, and breeding systems is included. In addition, information on plant environment and interaction, such as indicator status; better-than-state-level distributional data; elevation data; pollinators; herbivores; and forage, crop, and horticultural value, would further increase usability.

- Other aspects of facilitating use of both the published Flora and the computerized database include standardization, simplicity, and facility of access to information.

- Applications of a current continental Flora include direct use of descriptive information, use of information in a predictive manner, and use of information to make a variety of applied and academic correlations. All of these uses will promote development of testable hypotheses and create new directions for research involving botany as well as other branches of science.

- New information can be gathered and presented when the extended FNA database is able to interface with bibliographic and correlative research with other databases such as physiology, phytochemistry, etc.

- Information from the FNA database may become available either as subsets of FNA published volumes (i.e. for identification only, limited geographic areas), or as correlations of particular fields, sometimes involving data from other databases. Information of various kinds should be made available as hard copy and electronically as need arises.

APPENDIX 5

Report of Group 5. How is floristic information on characteristics and distributions used in phylogeny and biogeography?

Group Members

Bouchard, A.
Crosby, M. R.
Duncan, T.
Estes, J.
George, A.
Hartman, R.
Haynes, R.*
Kiger, R. (recorder)*
Miller, N.
Nevling, L.*
Wilken, D. (leader)*

I. INTRODUCTION

Phylogeny is the study of evolutionary relationships among species and higher taxonomic levels of organization. Phylogenetic relationships provide valuable information on patterns of inferred genetic similarity in time and space. Biogeography is the study of plant distributions in time and space. Biogeographical patterns provide valuable information at a number of levels of ecological organization and often are integrated with phylogenetic studies of particular plant groups. Phylogenies are based primarily on analysis of structural features of plants, whereas biogeographical distributions are inferred from geographical and ecological data that are tied to names of taxa. Both phylogeny and biogeography can be instrumental as predictive tools in other disciplines of biology. Both phylogeny and biogeography depend on the quantity and quality of data.

II. USE OF FLORISTIC DATA IN PHYLOGENY AND BIOGEOGRAPHY

Sources of floristic data include published local and regional texts that summarize taxonomic relationships, provide discussions of taxa, and give brief geographical and ecological data pertinent to the taxa of the area covered. Floristic data also can be obtained from computerized files established at regional herbaria and information centers, such as those operated by The Nature Conservancy, state natural resource conservation centers, and other federal and state agencies concerned with natural resources. Floristic data can be extracted from scientific articles; however, the diversity of literature with pertinent information requires considerable effort in retrieval and extraction. Data from such sources are often pertinent only to the variation or distribution within the region covered. Consequently, such data may not represent the full range of information for a particular taxon, a particular plant community, or a particular climatic or physiographic region.

---

* Asterisks indicate members of the writing team for the group

Major weaknesses of currently used floristic sources are:

- Failure to provide data on the full extent of variation in a particular taxon over its entire geographical range
- Failure to provide data on the full extent of species diversity, species richness, and other estimates of biogeographical interest for a physiographic, climatic, or geographical province
- Lack of parallel data for all taxa and all subdivisions of a particular region as a result of incomplete studies or knowledge

Types of data used in phylogeny and biogeography can be summarized as follows:

A. Phylogeny (systematic relationships)

- Patterns of structural (morphological, anatomical) variation
  - Vegetative features, including habit, stems, leaves, roots, and general life forms
  - Reproductive features, including floral parts (sepals, petals, stamens, carpels, and mature fruits) or spore structures
  - Microscopic features, including microscopic investigation of spores, pollen, seeds, surface features, and internal anatomy
- Patterns of phenology and reproductive biology
  - Duration of flowering and fruiting
  - Type of breeding system, including the type and amount of compatibility and the primary means of pollination or gene flow
  - Types of reproduction other than afforded by sexual systems and including both vegetative propagules and asexual systems involving the reproductive apparatus
- Patterns of chemical variation
- Chromosomal data at various levels of organization and observation

B. Biogeography

- Patterns of physiological variation and ecological distributions
  - Types of photosynthetic pathway
  - Tolerance of specific elements, pH, other features of soil chemistry, and soil structure
  - Environmental limits to rates of respiration and photosynthesis
- Fossil Evidence

- o Interactions with other plants and animals in the same region or in time
  - Pollinators
  - Herbivores, including dispersal agents
  - Data on plant communities, including cover, frequency, and associated taxa
  - Data on animal distributions
- o Interactions with the abiotic environment
  - Geological formations with an emphasis on abundance, surface structure, and chemistry
  - Climatic data, including variation in temperature, precipitation, their interactions and relationships through time

Both phylogeny and biogeography use various approaches to analyzing data, some of which are shared. Taxonomic relationships can be based on similarity of structure and function but can include objectively defined inferences of historical relationships tied together by ancestry and descent. Consequently, phylogenies are dependent on a synthesis of data derived from as many sources as possible. Biogeography involves the analysis of both taxa and their geographic distribution. From an ecological perspective, geography integrates the distribution of taxa with all aspects of the environment, including biotic and abiotic features. Biogeography also addresses the important questions of diversity, areas of endemism and origin, repetitive patterns, processes of migration, and anthropogenic changes. Both phylogeny and biogeography interact to provide useful insights on regional and local floras.

An integrated understanding of both disciplines can be used effectively in clarifying complex problems pertaining to the biota of a particular region. Such understanding can be extended and can contribute significantly to patterns and principles of biological diversity at a larger scale of geographical distribution. Detailed knowledge of biological diversity will contribute to management decisions regarding landscape ecology and conservation for both intrinsic value and potentially useful natural resources.

III. REQUIREMENTS TO MAKE FLORISTIC INFORMATION MORE USEFUL

Several critically important improvements in this area are needed:

A. A consistent, complete organization of characteristics that are parallel over all levels of taxonomic hierarchy is needed. Such an organization will benefit local and regional inventories and resource conservation agencies.

B. A factually reliable database that can be accessed in an easy and timely manner by a broad range of user groups, including researchers, teachers, students, resource managers, and lay biologists is needed.

These two solutions can be achieved in several ways:

- o Identify gaps in knowledge and implement field and laboratory research to provide necessary information
- o Establish rigorous standards for acquiring and organizing data
- o Establish a consistent approach to verification of data and avoid redundancy among separate but interactive data banks
- o Establish means and approaches to obtain more accurate data with particular attention to scale in both geographical and ecological distribution

C. Databases of continental scope pertaining to other major biotic and abiotic factors are necessary and useful in both phylogenetic and biogeographical analysis. The integration of these databases is considered to be important to the success of FNA and should be carefully planned and implemented in several ways. Aspects requiring consideration include improved hardware, software, and communications.

With current technologies, FNA should be expected to address issues other than just those related to monographic treatments of the vascular plants of North America. FNA can serve as a stimulus to future research on the North American flora if these issues are included during the original project, either as an integral part of the project or peripherally, in which FNA just ensures the gathering of the data. We were asked to examine such areas.

The responses can be grouped into the following five general categories listed below. Very little discussion occurred on each.

- o Computerization
  - Develop a specimen-based database
  - Develop databases for electronic identification, especially multiple-entry keys (polyclaves)
  - Develop a Geographic Information System, including a nationwide grid system
  - Provide a bibliographic database
  - Develop expert systems
  - Interact with remote sensing data
  - Identify gaps in the data
  - Make the data thorough, complete, parallel
  - Network with connections and nodes to avoid overload by users and/or contributors
    - . Allow rapid data transfer to peripheral data centers
    - . Connect taxon-based, specimen-based, and other databases
    - . Take advantage of databases from regional or state floristic projects

- o Ecological Data
  - Determine and designate indicator species for vegetation types, geologic formations, soil types
  - Identify critical habitats based on species preferences
  - Develop biogeographical criteria, such as uniqueness or representativity, for identification and selection of natural areas to be protected
  - Analyze abiotic and biotic variables to find meaningful groupings
  - Monitor rate of introductions based on chronological and distributional data
- o Expand concept of adequate morphological description and standardization
- o Act as an archive for data used in preparation of FNA treatments and for specimen-based data
- o Interactions with other groups
  - Other botanists
    - . Vascular plant systematists from other American countries
    - . Non-vascular plant systematists--North American and otherwise
    - . Botanists other than systematists
  - General public, especially amateur botanists

In summary, FNA is in a position to stimulate future botanical research by indicating gaps in our knowledge of the North American flora. We should carefully design our databases to gather data that can be used for such projects during the original data gathering process.

IV. NEW APPLICATIONS

- o Specimen data should be included in the databases and linked to taxon data.
- o A gazetteer of collecting localities should be developed and linked to other database files. This gazetteer should be updated and should include as precise locality data as possible.
- o Institutions with herbaria of important collectors should develop and publish itineraries of these collectors so that botanists studying their specimens will have guides to use when assigning locality data.
- o Committees including both systematists and non-systematists should develop standardized data categories that should be used for all taxa. As these data become available, they may be added to the database. It should be emphasized that all categories should be addressed for all species.

- Collectors and reproductive biologists should take care to standardize phenological and pollination data, including time of observation.

- Data from type specimens should be included in taxon data files.

- Development and adoption of some standard biological grid system for North America should be promoted. Included should be some provision for transforming all county/parish distribution data to that grid system.

Phylogenetic and biogeographic studies require access to comprehensive bodies of accurate floristic information, together with related data from many other disciplines. The printed Flora has been the traditional vehicle for presenting such floristic data, but its necessarily linear organization has limited the ways in which the information it contains can be accessed and viewed. Research in both phylogeny and biogeography focuses on the complex interrelationships among the many biological and physical factors involved in the evolution and distribution of organisms, and floristic information will be fully useful in such studies only if its presentation allows variable access to and flexible manipulation of data by categories, while maintaining overall integration.

Modern computer technology enables virtually unlimited flexibility in delivering, accessing, and viewing complex bodies of information, such as floristic databases. The core database that FNA builds as it prepares the contents of the printed flora can be utilized also for maximum exploitation of the floristic information in phylogenetic and biogeographic research.

We recommend that FNA undertake to provide, as early as practicable, the following diverse additional means of accessing North American floristic information in the core database and adjunct facilities:

- Downloaded database with accompanying user-friendly interface software on compact disc or equivalent, for local stand-alone microcomputer access and manipulation

- On-line network and telephone access to database through user-friendly interface, perhaps serviced from several regional facilities

- Searches, sorts, and tailored printed outputs, on a query/batch-process basis

- Polyclave (multiple-entry key) feature as part of the database interface for identifications or browsing by random-sequence characters

- Adjunct image-banking, for high resolution video display of illustrations and photos along with alphanumeric data from the database

- Linkage of database with GIS facility, for overlay mapping of floristic parameters

V. RECOMMENDATIONS AND CONCLUSIONS

1. A wide range of characteristics pertinent to the biology of plants and their geographical distribution is essential to the study of phylogeny and biogeography. Much information is available from museum collections, published literature and some electronic databases. However, gaps in knowledge at all levels of biological organization remain to be filled. Further studies and integration of available resources will reveal other directions of research.

2. Data used in phylogeny and biogeography can be supplemented in several major ways. These include (a) consistent, complete organization of data that are parallel over all levels of the taxonomic hierarchy; (b) a factually reliable database that can be accessible in an easy and timely manner; and (c) continent-level databases pertaining to other major biotic and abiotic factors important to phylogeny and biogeography and integrated with the FNA database.

3. The FNA project may stimulate new avenues of research activity in phylogeny and biogeography and may provide a focus for integrating ideas on a continental level. These may include (a) understanding of patterns of evolutionary change at the continental scale with regard to geological and climatological factors; (b) predicting floristic changes resulting from short-term changes in local or regional climate; (c) careful planning and success in conserving and preserving the indigenous flora and the diversity of plant associations at all levels of the ecological hierarchy; and (d) stimulating of interest on the part of the lay public particularly in the areas of conservation and also with regard to personal involvement in continuing study.

4. The availability and presentation of data pertaining to phylogeny and biogeography can and should remain flexible and versatile. This can be accomplished by means of traditional published treatments of data or syntheses. However, new approaches employing electronic documentation systems should be made available such that a diversity of user groups can access the data bank for periodic revisions based on new data or new analyses based on evolutionary principles.

5. Particular attention should be given to new approaches in phylogeny with regard to both acquisition and interpretation of molecular genetic data and its integration with currently used information and biogeographical implications.

APPENDIX 6

Report of Group 6. Resources: What workforce is needed and available to gather the floristic information targeted as desirable? How can the workforce be mobilized and managed? What existing resources could be networked with the FNA database?

Group Members

Argus, G. W. (recorder)*
Barkley, T. M.
Brouillet, L.
Gibbs Russell, G. E.
Gibson, A.*
Nelson, C. J.
Richards, J. H. (leader)*
Simpson, B. B.
Spellenberg, R.
Wilken, D.

I. INTRODUCTION

The Flora of North America project is seen to consist of three elements. The published volumes of the Flora have an associated, taxon-based database of the primary floristic information (Database I). Database I will develop into a more extensive database (Database II) that holds updated information and additional kinds of data and that is capable of interacting with other databases (Figure 1). We discussed the original set of questions separately for the published Flora, Database I, and Database II. The question of database networking is dealt with more extensively by Group 8.

II. SUMMARY OF GROUP CONCLUSIONS

A. *Information Targeted.* The information necessary for the published volumes and Database I is data for each taxon on morphology, phenology, toxicity, relative frequency, economic importance, habitat, and distribution. Correct names and synonymy are fundamental. References are extremely important, as are clear identifications of gaps in the data. Much of this data is already included in the format for the Flora. Our group recommends that notes on altitude (elevational data) be included with species distributions. Because many users need information about species relationships, authors should be encouraged to indicate species affinities.

Information that should be treated in Database II includes updates of information in Database I, expansion in detail of data presented in Database I, as well as additional kinds of information, such as data on pollination, breeding systems, pathogens, diseases, variability in individuals and populations, anatomy,

---

* Asterisks indicate members of the writing team for the group

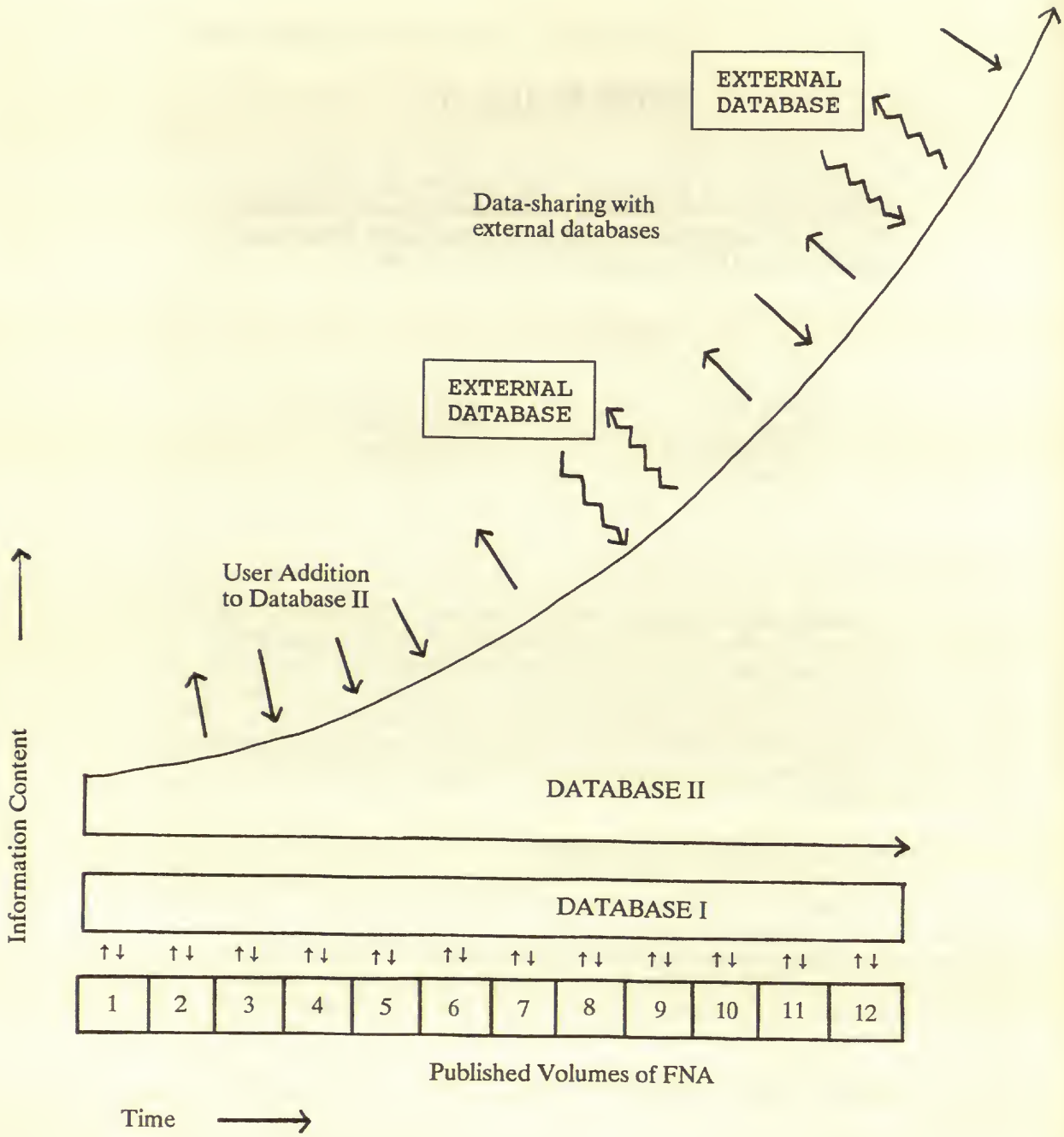

Figure 1. Growth of the Flora of North America database.

soil indicator species, and life history. Database II must be capable of interacting with external databases. These external databases will include specimen-level databases that contain detailed distributional information. Such distributional data may be correlated with geological, climatic, and soils databases. Similarly, external databases can provide a variety of data on the anatomy of specialized groups, rare and endangered species status, pollen/spore morphology, paleobotanical data, gene structure, chemistry, etc.

B. *Workforce for published volumes of the Flora.* The workforce needed and available to produce the published volumes in a timely fashion are the Convening Editor, supported by an administrative staff, the Editorial Committee, and the contributing authors. The Editor, staff, and Editorial Committee are already mobilized. Efforts need to be made, however, to ensure the cooperation of the contributing authors. We discussed the following four ways that participation may be encouraged:

1. Because publications are required for professional advancement, authors will benefit immediately through participation in and publication of their work in a prestigious Flora. The Editor should add to this benefit by sending letters of invitation to the authors' administrative superiors at the same time as to the authors, and letters of thanks to the same persons following completion of their contribution.
2. The Flora of North America, the Canadian Botanical Association/Association Botanique du Canada, and the American Society of Plant Taxonomists should pursue ways to give priority in publication in systematic journals to FNA monographs.
3. The National Science Foundation should be encouraged to give FNA-related projects additional priority in funding. Thus, the way in which a proposal contributes to FNA should be considered in evaluating proposals.
4. FNA should budget funds to provide authors with grants for travel, summer salaries, or sabbatical support. These grants will require additional funds from outside sources. We recommend that funding be made available through FNA to assist all contributors for several reasons. First, receiving a grant, no matter how small, will enhance the contributors' status within their institutions, thereby tangibly rewarding them for their work. Secondly, funds will enable some authors to travel to larger herbaria in order to study specimens and literature, thereby improving the quality of their work. Thirdly, funding will permit some authors to place themselves in an environment more suitable for concentrated floristic research and will therefore contribute to completing work on time. Finally, such grants will aid authors that do not have institutional support, such as retired botanists, to cover expenses that would otherwise be paid by employers. The ability to support authors financially will greatly expand the pool of knowledge on which FNA can draw and will enhance the prestige of FNA in the eyes of administrators.

Management of the FNA workforce is the responsibility of the Convening Editor, who has the final decision on any topic. We recommend that FNA hire a permanent post-doctoral level plant taxonomist to aid the Editor in covering groups for which a contributing author cannot be found and to take over treatments when an author fails to meet reasonable deadlines or to do a competent job.

C. *Workforce for Database I.* Database managers at the Missouri Botanical Garden are developing and maintaining Database I. We recommend that the Garden staff continue to be responsible for building this database. An additional staff person will be needed, however, to help enter data and process requests relating to this database.

D. *Workforce for Database II.* We predict that Database II development and use will grow exponentially throughout the publication of the Flora. This growth will result as Database II extends the information in Database I following publication of each Flora volume. Database II will contain additional data contributed by people specifically for the purposes of developing this database. Information accessed from external databases will be added to Database II only with permission of the owners of the external databases and as seems appropriate to the structure and function of Database II.

The structure and development of Database II should parallel and relate to that of the Flora. Its development, however, should not be the responsibility of the Flora Editorial Committee and staff, nor, necessarily, of Database I data managers. Database II should have a separate Editorial Committee and staff. A senior data manager should be hired to organize, develop, and coordinate this database. This person will help to obtain additional funds for the database and to integrate this database with other databases.

E. *Users' Group.* We recommend that FNA establish a Users' Group that will function in a support and advisory capacity to the Flora Editorial Committee and the Database II Editorial Committee. This group should be composed of people from botanically related fields, such as agronomy, forestry science, range science, from amateur groups such as native plant societies, and from other groups interested in the project. The purpose of the Users' Group is to advise the Committees on the needs of the users and, reciprocally, to publicize and support the Flora and Database II work in these related areas.

F. *Data Gathering.* Levels of data gathering could include taxon-based, specimen-based, and living populations or individuals. One charge of our group was to decide how to approach each level. The consensus of the group was that the Flora and Database I are taxon-based endeavors. Database II has the potential to accommodate all three types of data. Specifically, Database II should be able to interact with specimen-based databases. The development of such databases by individual herbaria would be extremely useful and should be encouraged.

The group considered, in addition, the question of potential sources of the three types of data. Sources for taxon-based data include published literature (previous floras, monographs, publications in other fields, etc.), unpublished literature (master's theses, archival notes, National and State Park files), and local societies. Specimen-based data are derived from collections, such as herbaria and seed collections, from collectors' notes, from Floras and revisions in which exsiccatae are cited, and from specimen databases. Data on individuals and populations come from published and unpublished literature, collectors' notes, records of arboretum accessions and botanical garden archives, and local agencies and societies, such as agricultural extension agencies and native plant societies.

III. OVERALL CONCLUSIONS AND RECOMMENDATIONS

A. Two linked databases should be created for the storage of FNA data. Database I should contain all the information needed to publish each volume, as well as supplementary data. Database II will contain data from Database I plus updates, extensions, and new kinds of information.

B. Database II should be coordinated with Database I, but it should be developed and managed by a separate staff. A database manager to organize and plan this effort should be appointed in the near future.

C. The mobilization of the workforce of contributing authors would be facilitated by grant awards made by FNA. Efforts should be made to obtain outside funding for such awards.

D. A Users' Group should be created to advise the Editors, Editorial Committee, and database managers about the needs and concerns of the users as the project progresses.

IV. SUGGESTIONS FOR FURTHER INVESTIGATION

Specimen Databases. These can provide a variety of collection, distribution, phenological, and habitat information to a taxon database, and are essential to supply fine-scale, accurately determined and vouchered distribution data for each taxon. Such distribution data will be required for many correlations with other data available in the FNA databases and will be fundamental in interaction of the databases with Geographical Information Systems. In order to obtain detailed taxon distributions necessary for these purposes, Database II must address the problem of computerization of extensive herbarium specimen records. It is estimated that there are about 60,000,000 herbarium specimens in the United States (P. Holmgren, pers. comm., 1989). A number of specimen databases are already in existence, especially in smaller institutions. A way must be found to assemble the contents of existing databases, as well as specimen data to be computerized in the future, in such a way that the distributions can be established at a desired scale of resolution both geographically and at any level of the taxonomic hierarchy. Problems relating to specimen-based databases will have to be solved in conjunction with the institutions involved because they have or intend to establish specimen databases. These problems include the following:

1. Determining whether Database II will itself contain a specimen database or whether it will serve as an interface to accumulate specimen data from a number of linked specimen databases maintained at a number of institutions and correlate this data with other data held in Database II.
2. Setting the data fields for herbarium specimen data that will be required by Database II.
3. Adopting a consistent method of quoting distribution records. The present "county" system is inadequate for correlations of distribution data to other data elements. A grid or point system should be adopted after investigating what other national standards or conventions are in use by other disciplines. Mechanisms to translate existing specimen distribution data to the new method must be found.

This possibly may include production of a gazetteer and use of computerized gazetteer data to convert conventional text data recorded on labels through the computer. (NB: This method was used by the PRECIS system with considerable success to convert to the quarter-degree grid system.)

4. Finding a means to ensure correct identifications of the specimens to be accessed by Database II.

5. Ensuring that specimen databases and herbarium collections are curated in parallel, so that specimens do not have one determination in the computer and another in the herbarium, or that duplicates do not have different determinations in different herbaria.

6. Determining how to deal with duplicate specimens in Database II. Where several databases hold duplicates of the same specimen, reporting these through Database II may not only be wasteful of computer resources, but could also bias reports and correlations by reporting the same specimen several times. On the other hand, it would be useful to use the FNA Database II as a means of recording and searching for the location of duplicate specimens.

7. Resolving implications for day-to-day curatorial procedures in herbaria whose specimen databases contribute to the FNA Database II. Specimen-based systems tend to fail for reasons of management rather than for inadequate computerization procedures. Because practices and staff duties in a number of institutions may be affected, this issue must be sensitively handled by the Database II manager and committee.

8. Determining how to use specimen systems from a number of institutions in such a way that the data can be provided without threatening the previous and future institutional investment in the systems.

9. Data transfer formats. The Taxonomic Databases Working Group (TDWG) represents a number of major taxonomic databases throughout the world. FNA must be represented at TDWG meetings, both to have a voice in setting standards and to ensure that it adheres to these standards. This is especially important because Databases I and II will doubtless influence the structure of other databases in North America not represented at TDWG. Through TDWG, FNA should especially encourage and participate in the development of data transfer formats so that information can be freely moved between systems. FNA must also look at transfer formats between databases of other disciplines that will be important for Database II.

APPENDIX 7

Report of Group 7. Computer Resources: What storage capacity will be needed for different levels of information? What software is available or needed for applications identified as desirable? What are considerations of access (ease, ownership, etc.)? How can we go about making information optimally transferrable and what software and hardware can we anticipate users having? How should networking be handled, and should permanent connections be established?

Group Members

Brigham, W. U.*
Duncan, T.*
Duvick, D. N.
Giles, R. H.*
Graham, R. L.
Gunn, C. R.
Haynes, R.
Johnston, M. C. (recorder)*
Kartesz, J.*
Macrander, A. M. (leader)*
Magill, R. E.*
Morain, S. A.
Morse, L.*

I. INTRODUCTION

The Computer Resources workgroup was charged with addressing technical aspects of developing databases both for support of manuscript preparation and for fulfillment of additional informational and analytical needs that may not be derived directly from FNA's published volumes.

We reached a consensus that these issues could not definitively, or even effectively, be addressed because many of the decisions that remain to be made concerning the ultimate function and scope of the FNA database(s) significantly affected our considerations. There are also many unknowns such as (a) the state of computer technology in the next decade; (b) the ultimate institutional status of FNA, i.e., whether it will constitute a free-standing institution in its own right or will be subsumed or fragmented among several institutions, e.g., a National Biological Survey; and (c) the mandate that FNA has or may assume in the establishment and management of botanical databases beyond the immediately defined goals of a published multivolume Flora. This document, therefore, will be designed to provide some general guidelines for addressing these five issues and to offer some specific recommendations for providing a planning group that will be equipped to deal with them as both the project and technology evolve.

There was recognition that the dichotomy between the database that will be derived directly from and/or for the preparation of manuscripts and the much larger and

---

* Asterisks indicate members of the writing team for the group

diverse database(s) that may be developed to support the complex research, curatorial, and resource management needs of the broad botanical community must be addressed immediately.

II. DATA-UNIVERSE

In approaching the establishment of floristic or more broadly defined botanical databases one might define the data-universe as including four major categories or bodies: (1) taxon, (2) bibliographic, (3) specimen, and (4) spatial. While these are interrelated and interdependent, each presents specific considerations for capture, maintenance, and dissemination. Taxon data include nomenclature with synonymy, descriptional information (at various gross external levels, anatomical and morphological levels, fine structure, and chemistry, etc.), natural history (e.g., phenology, ecology, interactions with other organisms), and identification aids such as keys and graphics.

Bibliographic data constitute a self-explanatory category.

Specimen data are essentially label-data and include identificational, locational, descriptive, and habitat definitions linked to individual collections.

Spatial data may be derived from specimen data or more general distributional information such as political unit (county, state, province) occurrence, or geobotanically defined units. The primary difference between specimen and spatial data is the nature of the management technique. Spatial data have the capacity to be accurately overlain with any number of ancillary data-sets or expressed in any number of configurations. This capacity may serve to increase both understanding and utility of floristic data.

It should be stressed that this definition of the floristic data universe does not necessarily define the universe of FNA databases. There are a number of capabilities and responsibilities that may be best addressed by other institutions or endeavors that may benefit and/or benefit from the database activities of FNA.

III. DATABASE POLICY AND STRATEGY COMMITTEE

The group recommends formation of a database policy and strategy committee. The committee is envisioned as being of parallel importance to the editorial committee. It will report to the head of the FNA project. A full-time hands-on database manager will be needed to implement the work.

The composition of this committee should include persons with the following expertise:

- A floristic botanist with FNA ties, i.e., one who is thoroughly familiar with defined goals and implemented procedures of FNA
- A member-at-large botanist, i.e., one who represents a user or customer of the FNA product
- A person experienced and expert in traditional tabular database design and management

- A person experienced and expert in geographical information system use and management
- A person experienced and expert in communications and networking via hard-wired lines, fiberoptics, and satellites
- A person experienced and expert in hardware design and performance
- A person experienced and expert in software design, programming, and debugging

The committee should be charged with the following somewhat overlapping issues for consideration:

1. Understanding and helping to define the scope and objectives of the database management strategy for FNA, particularly addressing what is to be included in the database and who will be the users of the database. Related issues will include the elements, fields, and relations of these databases.
2. Determining how FNA will insure the integrity, stability, and security of the FNA databases.
3. Deciding who will have access to the FNA databases and under what conditions; what protocols these persons will follow to read data or obtain subsets or reports; who will update, edit, and correct the data, and under what protocols.
4. Determining how quality of data can be assured or maximized, and what tests will be administered to data before entry to the final database.
5. Determining what the costs are for development, implementation, and maintenance over the term of the FNA project; what hardware and software the project requires; at regular intervals, perhaps annually, what new hardware/software FNA should add to its repertoire or convert to meet its goals and meet the needs of its users; what costs can be projected annually for inclusion in proposals for further funding for hardware, software, and database update.

IV. CONCERNS ABOUT THE FOCUS OF FNA's DATABASE DEVELOPMENT ACTIVITIES

A fundamental problem facing FNA is the need for attention to and definition of databases of use to FNA. Current computer applications have dealt primarily with the production of published volumes. Although the computer-assisted product is both computer-produced text and a relational database, long-term design of databases, questions of portability to new hardware systems, application of emerging database design standards, and problems of communication among databases external to FNA still need to be addressed.

Under the aegis of a database design and strategy committee, FNA should first identify existing academic, commercial, and private databases that contain significant data relevant to the identification, distribution, and related biological data on North American plants. After the survey is completed, the databases that FNA will maintain should then be defined. These databases should complement and not duplicate existing databases.

It is clear that FNA databases cannot (nor should they) contain all data relevant to North American plants, given the limited resources, the desire to publish 12

volumes in 12 years, and the existence of significant databases external to FNA. Therefore, FNA needs to clearly define its role in this context, develop databases that can be completed yet do not duplicate existing efforts, provide for maintenance of these databases indefinitely and define strategies to move these databases to new hardware easily. Furthermore, when appropriate, FNA should establish collaborative relationships with existing database projects so that mutual exchanges of information can enhance both the efforts of FNA and the projects currently under way.

With the several existing database projects currently under way, FNA may not necessarily be the appropriate umbrella organization to oversee the coordination of all botanical data. Further discussions are needed with organizations outside FNA to determine the role of the FNA databases in relation to these other activities. A set of published volumes is a worthy goal, and the future of FNA rests on its database management strategy. To be successful, FNA must have a clear, consistent focus and an achievable plan.

V. REPORT ON HARDWARE AND SOFTWARE

A. *Philosophy*. This subsection is based upon several considerations that are summarized as follows: first, we should work from a broad base resting on industry-standard, state-of-the-art technology. Second, however, we recognize that development of new computer hardware and software is taking place at a rate so rapid that we are not even sure of where we are on the curve (other than that it is going up). We should make decisions now that should not "paint us into a corner," but allow us to incorporate this new technology as it develops. Third, we must recognize that new technology does not instantly become dispersed among the user community. We must accommodate access from old computers which have migrated to home offices.

B. *Communications: on-line, dial-in*. This is the likely route for most direct-access users. The system should provide free or nominal cost telecommunication (we recommend "800 number" toll-free access with charges borne by the system). Modems should accommodate data transfer rates from 300 to 2400 baud.

C. *Networks*. Networks offer low-cost, high-speed access to computer systems by remote users. Some, such as BITNET, reach a worldwide audience. The host computers of the FNA database must be nodes on such a network. The FNA staff should explore and, if appropriate, encourage the use of NSFNET. Users desiring network access to the database would be responsible for establishing their own connections to the same or a compatible network.

D. *Other*. We have assumed that the ultimate FNA database will be distributed among several host computers. Certain search strategies may be difficult, if not impossible, on files residing on several hosts via conventional networks, such as BITNET. It may be necessary to provide a dedicated link among the hosts, perhaps via optical fiber or satellite to facilitate such queries.

E. *Off-line*. There will be instances where hard copy or transfer in electronic media will be requested by the user. The FNA database should be able to provide this output in a variety of formats, including printed text and tabular material, plots, diskettes, optical disks, and/or magnetic tape.

F. *Computer hardware*. We feel that it is inappropriate at this time to specify any particular computer system. However, it is advisable to identify characteristics

desirable in such a system. These include large multi-user environment, no restrictions in growth (or at least upward compatibility), multi-operating-system environments, and all major programming languages available.

G. *Peripherals*. Here again, it is inappropriate at this time to specify particular peripherals for either the FNA database or its user community. Some guidance in selection, however, is appropriate. First, it should be recognized that some specialized software includes device drivers that limit the choices of available peripherals. As an example, some GIS plotting packages have drivers only for specific plotters. Thus, a software decision should precede selection of peripherals.

A second guideline would be to chose general-purpose equipment rather than special-purpose equipment wherever possible. This is often an effective method for keeping peripherals up with current technology. As an example, a plotter is special-purpose equipment. It plots graphic files and nothing else. It is not possible to generate similar products of the same quality on other devices. As another example, a graphics terminal is also a special-purpose device. Here, however, it is possible to duplicate that function by running terminal emulation software on a micro-computer. The microcomputer is a general-purpose device. As graphics terminal technology changes, the terminal emulator can be upgraded at low cost to take advantage of that technology.

H. *Software*. Software presents a particularly difficult problem since large numbers of packages are in current use. The FNA database should be developed around the idea of portability with future system enhancements, as well as data transfer, in mind.

A single software package for all parts of the ultimate FNA database may not be possible or even desirable. All future development of the various components, however, should move toward a common strategy of data structure to facilitate interaction among the various parts. FNA should investigate data structures of existing databases in other disciplines to facilitate exchange of information. A common data strategy would eliminate the need for most conversion programs and allow more rapid transfer of data between the various parts of the ultimate FNA database.

VI. THE TWO-LEVEL DATABASE CONCEPT AND FUNCTION

As the editorial committee for published volumes of FNA develops databases from which to produce texts for Oxford University Press's yearly publication of a volume, they should become knowledgeable about the status of operating and planned national and international biotic and abiotic databases. Before the first volume is published, the committee should start to build links to these databases on a selected basis, forming a data transfer network. Abiotic databases such a weather, geological survey, and soils would be entered using map coordinates. Biotic databases such as plant surveys, herbarium inventories, and nomenclature could be extended using scientific names.

Most of these data from this data network would not go into the published volumes. These links permit greater flexibility and widest user applications of the networked databases through the exchange of comprehensive datasets.

In the short run the printed volumes will be FNA's most valuable product, but in the long run the networked databanks will be the most valuable product to the widest number of users.

The editorial committee might best reach this network goal through the recommended Database Policy and Strategy Committee, whose primary goal would be to advise FNA on hardware, software, and data portability.

A. *Floristic database*. There is to be created a computer database associated with the 12 volumes (hereinafter called the Flora). That database will have many elements, e.g., species number, names, synonyms, technical description, chromosome number, province/state and county distribution, habitat, literature. These are a relatively few elements; and even though there are many thousand plant names, the elements for each plant are few. This initial database is illustrated in Figure 2(a).

To meet the expressed needs of botanists, the elements can be used together, e.g., to sort all plants in an area with similar chromosome numbers. There are definite uses of the database to be created to produce the Flora and to support the interests of users where interests are restricted to those elements in the database created in order to produce the Flora.

B. *Other databases*. Many other users have expressed interest in the Flora and its database. That database becomes useful only if other databases are available. Figure 2(b) shows the data but with added elements, factors that can be related for a variety of purposes, a variety limited only by human creativity. The number of uses is related to the potential permutations of the elements, i.e., n!.

There are many databases in existence that are clearly related to the elements in the Flora, for example census data, climatic data, soil characteristics, and wildlife that consume certain plants. There are other databases related to the Flora and the community concerned with and involved with plant taxonomy. These include 1) a herbarium location address and information system; 2) a specimen database (a within-herbarium collection system); 3) a synonym list; 4) a bibliography; 5) an authority list; and 6) a collectors list. Where to find samples of specimen X that can be obtained on loan is an obvious application of having the Flora, its databases, and access to a related database on herbaria. This expansion to include related databases is shown as in Figure 2(c).

Less intimately related are databases on plant substances, toxicity, photosynthetic rates, uses for human nutrition, and noxious weed potential. These need to be identified, means of access clearly identified, and ease of access (telephone linkage, costs, etc.) provided by the expanded database and its staff-created software. Certain software and aids need to be provided to enable easy use within the floristic community such as getting element 1 from source C, element 2 from source D, and element 3 from source G, combining them with two elements in the Flora system, and manipulating them with a statistical program (e.g., SAS) or creating new maps for specific users.

The new system adds elements where feasible and practical. In some cases, once files are collected and a new index created, that index should be stored as a new element, not discarded because it is "artificial" or because it was calculated and can be recalculated. Thus the system grows in included elements but, just as important, also grows in direct easy links to existing databases. The intent is to expedite and make more effective use of data in the floristic community. In some cases this will be done best by adding data elements to Flora; in others, it will be by linking to other elements.

Quality control will remain a problem. The user must be aware that the Flora staff may provide efforts to evaluate data, but decisions on use (and associated risks) will remain with the user.

The expanded and evolving system appears as in Figure 2(d) with the Flora completed or nearly complete, and the residual database shown in the solid box, and the enhanced linkages with a long list of dying and newly-created data sets being implied by the gray area.

The relations possible then become those of distributed data sets, those specific for regions, provinces, or states. These are the solid, self-contained area-specific parts of the central system shown in Figure 2(e). Each subunit may link to the national or international data sets as needed to meet its specific users objectives.

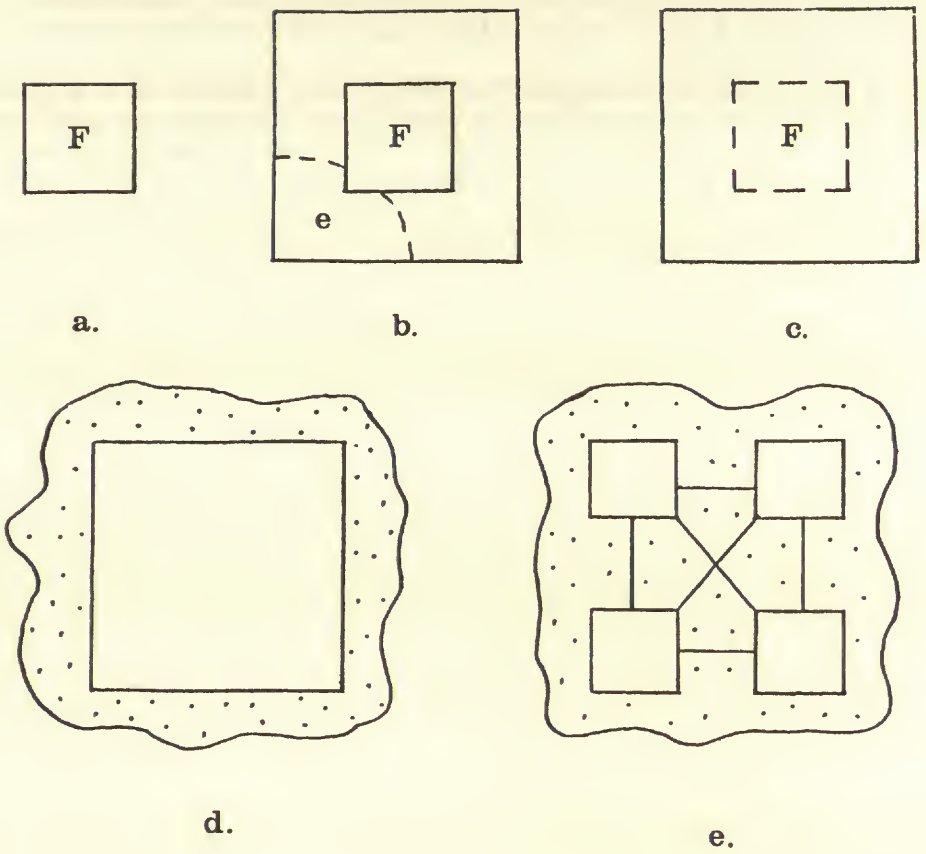

Figure 2. Development of Flora of North America-related databases. a) Initial FNA database; b) FNA database with added elements; c) expansion of FNA database to include related databases; d) residual FNA database with dying and newly-created data sets (stippled area); e) subunits of central FNA database showing linkages.

APPENDIX 8

Group 8. Interaction and collaboration with projects of other scope

Group Members

Beecher, C.*
Catling, P. M.
Fay, J.
George, A.*
Hartman, R. (recorder)*
Hickman, J. C.
Nevling, L.
Rhoads, A. (leader)*
Wagstaff, J.*
Young, D. A.

I. INTRODUCTION

For FNA to be successful in the shortest possible time, it is important to identify all relevant resources that its authors or users may need. Authors and users will have different requirements. At the outset there should be a focus on the most current taxonomic needs of the authors. But the ultimate scope should be as broad as can be conceived. For this reason, we include both sources of data sets of taxonomic and related information and sources of data sets of supplementary information.

Existing data sources and technologies should not be duplicated; rather, conduits for passing data must be established. The key to relating data from various sources must be the plant name. FNA should provide the standard, and it is hoped that other databases will find it in their interest to conform, perhaps through the use of synonymy tables.

There are many kinds and levels of data and databases. In our discussions, we have sometimes blurred the distinctions between them. For our purposes two major types of databases should be considered: those that are built of taxon-based or specimen-based primary data, and those which are compilations of derived or synthesized information.

Taxon-based and specimen-based databases are the fundamental workhorses of the FNA project because they provide access to raw data. It is best to recognize FNA as a database of synthesized data, because it will accumulate data from primary sources and organize it in a fashion that will be distinctly greater than the sum of its parts.

By recognizing that FNA is a database of synthetic knowledge we understand its unique abilities and its limitations. For instance, since it is reliant on primary literature, it is important from the outset that the authors have access to the primary sources.

---

* Asterisks indicate members of the writing team for the group

What this project needs to do immediately is adopt a first approximation of a complete list of accepted names. This list will form the working version of the state of the synthetic art. Currently existing databases can provide authors with a wide range of information. These authors will then base their treatments on the greatest possible availability of information.

II. TYPES OF INTERACTIONS WITH FNA

A. High priority information for authors.

Contributors to FNA will need access to basic data sets for preparing text. In some cases, these should be provided routinely; in others, advice should be given on the content and accessibility.

First priority must be the *Kartesz Checklist* (Kartesz and Kartesz, 1980). Each contributor should receive a subset of the taxa to be included in FNA to use as a starting point.

Essential for text preparation are bibliographic sources like *Index Kewensis* (1891--); the Gray Herbarium Cards; the *Draft Index of Author Abbreviations* compiled at the Royal Botanic Gardens, Kew (1980); *Botanico-Periodicum-Huntianum* (1968); and F. A. Stafleu and M.S. Cowan, *Taxonomic Literature*, second edition (1976--1988).

The *Australian Plant Name Index*, being compiled as a database by the Bureau of Flora and Fauna and the Conservation Commission of the Northern Territory, should be used for primary nomenclature and bibliographic data when appropriate.

Secondary checking of taxonomic nomenclature and distributional data will involve national and regional Floras and checklists. An index of these should be available for all contributors. That for Floras will be relatively easy to prepare. Checklists will be both published and unpublished; the latter will need considerable inquiry at the regional and local level to locate, and usefulness will depend on the reliability of the data (e.g., whether or not they are backed by voucher collections). Data on toxicity should be available from *NAPRALERT* and *Toxline* and the Poisonous Plant Information System at FDA. Accessibility to these should be given high priority. If conservation status is to be included as standard in FNA, then access must be provided to national and state lists of rare and endangered species.

Supplementary data that should be accessible if required by a contributor include chemical data, habitat files, cultivated Floras, and plant morphology, economic uses, ethnobotany, and chromosome number.

Both the contributors to the published FNA and users of the FNA database will need a description of the content of each data set, how it may be accessed, and how to obtain further information.

B. Monographic Work.

Monographic work should be viewed as a research avenue distinct from but closely related to the FNA project. It is much more detailed and comprehensive in scope and requires far more time than preparation of a Flora text. Published monographs will provide a source of information for Flora accounts. It is vital that, where a monograph is in progress, the Flora not be delayed in anticipation of its completion. This would prolong the Flora project unacceptably. It should not be taken for granted that a monographer will also be a Flora contributor. Some systematists consider their primary role to be that of a monographer and have no

interest in the more concise account. They should not unwillingly be pressed into service. If a monograph is in progress but not scheduled for completion in time for FNA, it may be possible to arrange for a Flora staff member to work in conjunction with the monographer, while being careful to avoid duplication.

C. Direct Incorporation of External Data Sets into the FNA System.

In some cases external data sets may contain information that can be transferred directly into the FNA system. Centralization of databases should be carefully considered in the case of rare information.

Issues that must be addressed are: reliability and volatility of data and feasibility of centralization of databases. All data to be directly included in FNA must be verifiable, that is, referenced by specific primary sources. Volatile data (data that changes with time) can be handled by periodic (yearly) updates from the external source.

D. Access to Supplementary Data Sets.

It may not be physically possible to centralize all relevant databases. Therefore FNA should provide information on availability and access to other information sources in the form of directories, lists, and annotated bibliographies, and names and telephone numbers of contact persons.

E. Relationship of FNA to Geographic Information Systems.

Another type of relationship should exist between FNA and sources of biotic and abiotic data about the environment. Information on soils, geology, topography, distribution of other types of organisms, cultural features, and climate are components of a Geographic Information System (GIS) of which floristic information is also a part.

FNA should be viewed as one of numerous modules that comprise a GIS. In order to maximize the applicability of FNA information in Geographic Information Systems, the issue of a grid or other geographic identification system should be addressed. The U.S. Geological Survey should be consulted.

F. Additional Considerations in Establishing Interactions.

A determined number of high-priority conduits should be established between database providers and the central committee of FNA. The conduits would then be available to authors, database providers, and others.

Supporting data sets could be housed centrally, perhaps at the Hunt Institute for Botanical Documentation or in an existing integrated database system (IDBS) such as SEEDIS operated by Dean Merrill at Lawrence Berkeley Livermore. The central database system should be inexpensive and easy to use. Perhaps an integrated database system could handle the taxon-based, specimen-based, bibliographic, and supplementary databases. If this is not feasible, then bibliographic and supplementary databases could be in a central on-line depository such as the Hunt Library and the rest in the IDBS. [The TROPICOS database system at Missouri Botanical Garden, which is being used currently for FNA, has these characteristics and can accommodate such databases. --Eds.]

Each provider of a database (chemistry, toxicology, etc.) would review the completeness of the database in the area to be covered prior to the deadline for receiving copy for each volume. In addition, a complete revised copy of each database should be sent to the central repository each year.

Each conduit should have the following characteristics:

- Data items desired by the central repository should be listed and defined.

- Recompense and recognition should be provided. The central repository needs to provide some kind of payment for products coming from the conduit--perhaps grants, co-authorship of publications, titles such as "consultant to FNA," "special assistant," etc. These should be stated.

- Key data elements with unique codes should be adopted at the level of specimen, organism (taxon), place (grid and county), chemical (CAS, NAPRALERT , etc.), and literature citation.

- Data format restrictions should be minimal. The data should be understandable by humans and should be in ASCII with record and field delimiters; the capability should be provided for future inclusion of abstracts, full text, and graphics (illustrations, maps, etc.) on-line.

- At first raw data should flow to the central repository. Then plans should be laid for more detailed organization of information to flow on the conduit.

III. RECOMMENDATIONS AND CONCLUSIONS

1. The list of project and data sources included as Table 7 (p. 83) in this report must be considered preliminary. In order to develop a more complete inventory, we suggest a mailing to workshop participants that would include the present list and ask for contact person(s), addresses, telephone numbers, as well as additional sources as a first step.
2. We suggest that FNA make contact with each project/source to a) inform them about FNA project, b) determine the relevance of these external sources of information to FNA, and c) initiate appropriate interaction. Symposia may be a useful mechanism for informing other groups about FNA.
3. We suggest that FNA provide a clearinghouse or coordinating function, providing direction on where additional information can be found; and that the FNA database include references to species-specific data sources, taxon by taxon, as well as general references to other data sets that can provide more information.
4. We suggest that FNA encourage the establishment of specimen-based databases on a state or regional level and the use of standardized data fields.
5. We suggest that the FNA Advisory Committee be reactivated for the purpose of guiding interactions with external projects and data sources. A title such as "Special Assistant to...." should be created to dignify Advisory Committee members and individual consultants.

6. Identify gatekeepers at each end of a conduit , e.g., Kartesz and maybe Young at the taxonomic end of the toxicity conduit, and Wagstaff at the toxicology end. Gatekeepers should be facilitators, not police--anyone who has need and is responsible should be allowed free access to a conduit.

V. CONCLUSION

There are many data sets that should be used in preparing the published FNA and the database. An initial list is included but it should be augmented as more people are consulted. Some are essential tools for contributors; others should be advertised as available for both contributors and users. An inventory of all data sets must be prepared. A feasibility study should be made to determine which sets should be gathered into a central bank and to set up means of communication with others.

References

Kartesz, J., and R. Kartesz. 1980. *A Synonymized Checklist of the Vascular Flora of the United States, Canada, and Greenland.* Chapel Hill: The University of North Carolina Press.

Index Kewensis; plantarum phanerogamarum nomina et synonyma omnium et specierum, a Linnaeo usque ad annum MDCCLXXXV ... 1893--95. Clarendon Press, Oxford. 1893--95.

-------- *Supplementum* ... 1901--. Clarendon Press, Oxford.

Meikle, R. D. 1980. Draft index of author abbreviations compiled at the Herbarium, Royal Botanic Gardens, Kew. Her Majesty's Stationery Office, Basildon.

Lawrence, G. H. M., A. F. G. Buchheim, G. S. Daniels, and H. Dolzel. 1968. Botanico-Periodicum-Huntianum. Hunt Institute for Botanical Documentation, Pittsburgh.

Stafleu, F., and R. S. Cowan. *Taxonomic Literature*, edition 2. Regnum Veg., vols. 94, 98, 105, 110, 112, 115, 116.

APPENDIX 9

Report of Group 9. What should the long-term management scheme of a project such as Flora of North America be?

Group Members

Crosby, M. R.*
Estes, J. R. (leader)*
Heywood, V. H.*
Kiger, R.
Kosztarab, M.*
Miller, N.*
Murray, D. (recorder)*
Shetler, S.*
Stein, B.*
Stuessy, T. F.

I. INTRODUCTION

This group attempted to provide suggestions for the long-term management of the Flora of North America project in terms of its relationship to potential national and global biological surveys; in terms of its eventual composition (e.g., should it be a center for taxonomic information, for all plant-related information, or something intermediate?); and in terms of the need for perpetual maintenance and continued development of the database.

II. SUMMARY OF GROUP CONCLUSIONS

We began our discussion by first assuming that FNA consists or will consist of the following components:

- o hard copy printed version of the Flora in 12 volumes written and produced in an electronic database environment and therefore accompanied by
- o a computerized database of essentially floristic information both leading to and derived from the writing project, which can provide the basis for establishing
- o an encyclopedic database with more kinds of data in greater detail and with links to other databases.

Given that structure of FNA, we then dealt with each component and each question separately.

Concerns were expressed that the management and editorial structure must be so designated as to ensure the consistency, accuracy, and momentum required to achieve a satisfactory printed Flora. There is both individual and corporate

---

* Asterisks indicate members of the writing team for the group

responsibility, but there must be an editor-in-chief or a chair of the editorial committee who will be the ultimate arbiter.

The editorial committee must consider very carefully the actual workload that will arise and remain for a minimum of 12 years. There will be the need for regular meetings to discuss progress, timetable, staffing, funding, and technical matters too numerous to mention but generally devoted to establishing additional norms and standards. All members of the editorial committee presently see every manuscript. It might be better to divide the labor or to involve only half the committee each year with Family responsibilities. There will be new manuscripts to see, older ones in various stages, and galley to proof--all simultaneously. It might also be useful to employ certain members of the editorial committee or to invite additional members for special assignments such as the checking of geographic ranges, chromosome numbers, etc. These special subject, or thematic, editors could develop very finely honed abilities that other reviewers might not have. Who will review the manuscripts for taxonomic consistency, that is, to see that there is a comparable level of generic splitting within families? How will the ensuing debates be resolved?

With respect to the limited database directly associated with the writing project (database I), what is the minimum level of documentation required for such topics as typification and geographic distribution? It is understood that each contributor must rely on the best of the earlier work, but it should also be clear what herbaria have been consulted, what type material for each species has been seen (holotype, isotype, syntype), and what specimens document geographic range that departs from what has been accepted. Specimen-based data should be provided in as much detail as possible, and it is suggested that for geographic distribution, authors should cite at least one specimen for each unit-area (level to be agreed). We must have the ability to know the sources (at the specimen level) for all elements in the database (at the taxon level).

A. We suggest that an advisory board be formed for database I with overlapping membership with the editorial committee. This board would review matters of hardware, software, and finances.

For the long-term management of the encyclopedic database (database II) we feel there must be a task force with membership from beyond FNA as it is now constituted. They must consider the scope of database II before we can propose how it might be managed. What other sources of information are there that database II could use by network or by wholesale incorporation?

B. Task force for the expanded database. The long-term management of FNA, in terms of an expanded database (or taxonomic information center), can only be addressed once it has been determined what its composition and structure will be, in the light of relationships with and developments in other cognate databases, information systems, or biological surveys. It is therefore necessary to set up a task force to identify these cognate databases, etc., and to explore the extent to which they can provide modules of information that can be networked or incorporated.

The Task Force should include a representative of the FNA committee and other appropriate organizations at the national and global level. These latter might include:

The National Biological Survey
The Biological Survey of Canada
The Nature Conservancy
The U.S. Department of Agriculture
International Union for Conservation of Nature and Natural Resources (IUCN), including the Threatened Plants Unit (TPU), Botanic Gardens Secretariat (BGCS).
United Nations Environmental Programme (UNEP)
Global Environmental Monitoring System (GEMS)
BIOSIS International
CAB International Abstracts (including Plant Breeding Abstracts, Forestry Abstracts, Weed Abstracts)
European Taxonomic Floristic Biosystematic Documentation Systems (ESFEDS)
International Legume Database and Information System (ILDIS)
MedChecklist (Geneva and Berlin)
NAPRALERT (Chicago)
Bureaux de la Flore et de la Faune (Paris)
EEC Commission (Brussels)
National Institute of Health (Washington)
Botanical Research Institute (Pretoria; houses PRECIS)
Taxonomic Databases Working Group (will become IUBS Commission in 1988)

It might be appropriate to convene a workshop of representatives of these and similar organizations to explore the possible links that could be established, identify the gaps, and recommend further action.

It has to be remembered that all the above organizations already have relevant databases that will presumably continue to expand and develop. The kind of questions that will have to be asked are:

- What fields are covered?
- What are the gaps?
- Is networking possible in this case?
- What are the financial and copyright implications?
- What are the technical problems of information exchange?
- What kind of corporate structure might be envisaged with participating partners contributing one or more modules?
- Is there a possibility of establishing a national/international clearinghouse for these databases?

Similar questions have recently been addressed by several of the world's leading conservation databases and associated organizations at a workshop held at IUCN General Assembly in Costa Rica (February 1988); the report of the workshop should be studied.

C. The position of FNA as a potential center, either for plant taxonomic information alone or for all plant-related information, depends on several outside factors, e.g., establishment of a national biodiversity survey. But it is clear that the long-term maintenance of either a center for plant taxonomic information or a center for all plant-related information will depend on public funding, since such centers cannot attract enough users' fees to be self-supporting or self-perpetuating.

Database I can certainly constitute at least the basic components of a center for plant taxonomic information. Taking database I as a given, a scheme for its long-term maintenance must include provisions for updating its basic content (morphological, distributional, nomenclatural, cytological, and bibliographic data), and provision for making updates available to consumers. Various ways of providing updates may be imagined, e.g., on-line queries, printouts, or diskettes; but since FNA has a contract for producing the initial book-form of the data, any means of providing updated information must be worked out in relation to this contract.

An ongoing program of updating database I under its original set of parameters or slightly expanded ones could be more or less passive (adding appropriate information as it became available from non-FNA projects) or active (along the lines of the revision of volume one of *Flora Europaea* now under way). Certainly a scheme for keeping the bibliography up to date could be established.

A case can be made for the position that all information about plants and much information not about plants is pertinent to plant taxonomy and thus database I could evolve into a center for all plant-related information; in practice this is probably not a viable position. Rather, database I is more likely to become a module of a center for all plant-related information, providing the basic taxonomic/floristic component of such a center.

D. In light of the above, for long-range funding, it is essential that FNA management make an effort to identify an umbrella organization such as the National Biological Survey or Center for Biodiversity Research, if and whenever the funds to support the objectives of these projects are available. H.R. Bill #4335, now in Congress, includes 25 million dollars of support for the first three years for basically a survey of North American biota. Obviously FNA fits well under such new legislation, and there could be the possibility of securing long-term funding from these sources, rather than depending on foundation grants that normally support short-term projects of one to five years.

Other possible sources of funding are the Biological Survey of Canada. It now supports research on terrestrial arthropods, but an expansion of their view could be sought. The Canadian government is well aware of the current threats to the biota, and it may well appreciate the need for the information FNA will provide.

III. OVERALL CONCLUSIONS AND RECOMMENDATIONS

The group considers that FNA should be viewed as: a) the 12-volume printed Flora, b) a cognate database, and c) an expanded encyclopedic database.

It proposes that the editorial and management structures for the hardcopy Flora need a certain amount of expansion and definition, particularly with the view to covering the specialist or thematic editing tasks. It recommends a division of labor

on an annual basis for the duties of Family editing. It recommends that the management of the cognate limited database have a separate structure, albeit linked with or a subdivision of the main editorial committee. For the expanded encyclopedic database it recommends the establishment of a task force to explore the availability of networking with relevant databases that could provide modules. It strongly urges FNA to explore the possibility of some federal agency or organization such as the proposed Center for Biological Diversity as an umbrella or parent organization so as to ensure the possibility of long-term, core funding.

IV. SUGGESTIONS FOR FURTHER INVESTIGATION

1) Transfer and exchange formats to facilitate communication and data exchange with other databases.

2) Protocols for data capture and handling by contributors to the Flora and its databases.

APPENDIX 10

Report of Group 10. Additional topics identified in the brainstorming session

Group Members

Bouchard, A.
Flint, H.
Hall, D.
Hayne, D. W.*
Larson, G.*
Leiser, A.*
MacBryde, B. *
Packer, J. G.*
Shultz, L. M.
Taylor, I. E. P.
Waggoner, G. (leader)*
Whetstone, R. D. (recorder)*

I. INTRODUCTION

The charge for Group 10 was to review all issues raised in the morning's general brainstorming session and to assure that all ideas were addressed. The Group decided that most of these brainstormed ideas would be covered in the deliberations and recommendations of Groups 6, 7, 8 and 9. To assure that all significant issues were treated, the Group considered and made recommendations on brainstormed ideas that were believed to be most significant and less apt to be considered by the other groups. The results of this session were categorized into four major areas: A. Published Flora, B. Database, C. Public Awareness, and D. Miscellaneous Suggestions.

II. SUMMARY OF GROUP CONCLUSIONS

A. Published Flora

Publication of FNA hard copy is the first priority. Several problems were discussed regarding the actual publication of FNA. Answers to these problems were not always supplied, but it was felt that the editorial board would need to address these issues. Issues included:

- o Gaps in existing database
  - How to fill geographic gaps in existing database (gaps such as those illustrated by the discussion of Gibbs Russell for South Africa)
  - How to set priorities for filling these gaps

---

* Asterisks indicate members of the writing team for the group

- How to obtain coverage for "orphan" or "cabbage patch" taxa, those taxa not covered by any specialist

  - Advertise and issue invitations to authors who would be willing to treat these taxa
  - Develop a program to recruit retired botanists and private collectors (or groups such as native plant societies)

Publication of ancillary volumes tailored for special user needs/tastes should follow publication of initial volumes. A first priority after publication of FNA should be a single-volume identification manual, or at most two volumes (Monocots and Dicots). These manuals would contain keys and very brief descriptions. An additional supplementary publication would be a specimen-based atlas.

The Guide for Contributors should include certain definitions, cautions, and information to promote uniformity.

- Escaped species:

  - Treatment of "exotics" might vary depending on the degree to which they have escaped--rare or occasional versus widespread escape. Precise definitions of terms used should be provided to contributors and other users.
  - Native escapes (for example, *Robinia pseudoacacia* L.) may vary from expanded range (in East and Central U.S.) to widespread introduction and escape in the Western U.S. (California, Oregon, and Washington).
  - The Guide for Contributors should identify experts on a regional basis to assist regional editors in determining escaped status.

- The Guide for Contributors should caution authors to treat "volatile" data with care and not publish them in FNA. Volatile data include premature determination of rare and endangered species that have not been thoroughly documented, and might also include nomenclatural and taxonomic changes that have not withstood the scrutiny of other taxonomists.

Throughout the preparation of FNA, authors and editors should be constantly aware of the various sectors of potential users and provide services (treatments) tailored to this audience.

B. Database

In considering database issues, we asked whether it was a better strategy in generating support for FNA to emphasize the publication of the volumes at the expense of work on the database, or whether the latter should receive greater emphasis now and be developed in parallel with the publications schedule.

Increasing emphasis on database development carries with it the danger of reducing the momentum in the publication schedule, especially if there is diversion of funds. Appreciation of this fact is general, reinforced by well known historical

examples and caution of experienced database developers. Yet we must acknowledge that a computer-accessible database of the FNA information would greatly increase the usefulness of that project to non-specialists, in making information on distribution easily available and in particular as a tool in plant identification by multiple entry keys. This last potential is now mostly unknown outside the field of a few specialists.

Further, having the FNA database linkable to other subject matter databases would greatly increase the potential usefulness of FNA in research in resource management, plant and animal ecology, geography, and numerous other fields. Unlinked to these other data files, the usefulness of the FNA database to the non-specialist seems to be limited to the determination of distribution and help in plant identification, important as these two tools may be.

Thus, with the objective of continuing and perhaps increasing support for FNA, the parallel development of the database seems essential because it would so greatly increase the usefulness of the FNA data and broaden the scientific support base for the project. A separate committee is needed within FNA for advising on database problems and for pushing the development of the database. Obviously, there are questions of fiscal support here. But if it is sound reasoning that the availability of the database will increase the scientific base of support, then it will be good strategy to make the database available, group by group, as the volumes of FNA are published. This strategy has the advantage that if the reasoning is not correct, the diversion of resources to the database can be reduced.

In the more technical questions about the database, there is general agreement that the FNA database should restrict its content to data from FNA, with a coverage somewhat expanded beyond that of the published volumes, but not including extensive coverage of environmental factors. The format should be dynamic, capable of accepting different data, if necessary, in the future. Future data entry should be screened for quality, probably by a specific committee.

There is an immediate problem with compatibility of other databases with that of FNA; perhaps this can be solved by computer programs, but if so, this point alone stresses the need for specialized personnel.

Obviously, even though the FNA database should not attempt to include massive files of environmental information, for example, it will be necessary in setting up the activity to plan carefully so that FNA can easily be linked to other data files.

C. Public Awareness

Our group recognized that in order to ensure success of the FNA project, there is a need to generate broad public awareness and foster support for the Flora. A great variety of methods was discussed toward achieving this goal:

- Those in attendance at the FNA Workshop should communicate the mission and benefits of the FNA to their academic communities and political bases back home. Targeted for contact should be academic and agency administrators, congressional delegations, and public and private funding organizations, all of whom could promote the idea of FNA and possibly provide access to financial resources.

- The existence of FNA should be advertised in one or more national publications. Of particular value would be articles in popular magazines such as National Wildlife, Audubon, or Natural History.

- The current mailing list for the FNA Newsletter should be expanded to include political bases, funding organizations, conservation groups, garden clubs, etc., especially those likely to be friendly toward FNA.

- An attractive, eye-catching logo for the FNA Newsletter and other public relations documents should be developed.

- Nonbotanical user groups (e.g., governmental agencies, planners, resource managers, researchers, landscape architects) who are likely to be supportive of FNA should be identified. The usefulness of FNA should be communicated to them and thereby justify the expenditure of funds toward the project. The educational and practical values of FNA should be emphasized.

- Regional (and taxon) editors should be used to help inform organizations about FNA. The editors should have access to a list of organizations and potential user groups who are likely to promote FNA in their region. Additionally, scientific editors for the electronic and printed media should be informed of the project by regional editors.

- Pre- and post-publication fliers for each printed volume of FNA should be developed.

- As a means of publicizing the need for FNA and research that would complement it, distribution maps should be published in the FNA Newsletter for taxa that are perceived to be in need of better documentation.

- Primary and secondary level educators should be informed of the values of the FNA product. Summer educational programs and workshops are good vehicles for accomplishing this. Audio-visual materials on FNA should be developed for distribution to educators.

- Public museums and botanical gardens throughout North America should be communicated with as a means of getting the word on FNA out to the most relevant audiences.

- A professional publicist who has experience in the natural resource field may need to be employed. This possibly could be done on a regional basis, with each regional editor in charge of publicist duties.

D. Ancillary and Miscellaneous Suggestions for the Editorial Committee

- There should be an ongoing update (report) on the status of discussion, current considerations, and future possibilities of the overall project.

Publish as Volume 2 (or more) of the proceedings of this workshop information on the history and development of FNA activities and publish on a regular basis documentation of the ongoing evolution of the project.

- FNA should provide a clearinghouse for "hidden" or obscure documentation--e.g., MS theses, state Department of Agriculture reports on weed escapes, etc.

III. OVERALL CONCLUSIONS AND RECOMMENDATIONS

Four major themes emerged as important concerns that we categorized as A. Published Flora, B. Database, C. Public Awareness, and D. Miscellaneous Suggestions. The primary Flora issue mandates publications as the top priority and primary objective offered. Ancillary volumes will become necessary as data gaps are filled, and these volumes should be tailored to specific user groups. Database issues include development of policy, which should define relationships of the database to the Flora design, etc.; to what extent databases or other resources are available and of primary importance for the project; documentation and verification of data entry; and access control to sensitive, copyrighted material, unpublished original research, etc. Public awareness issues are manifestly diverse. We should target: the general public to "sell" the need for FNA; the scientific community on importance and involvement; and the government about the benefits of FNA. Miscellaneous concerns of an immediate nature are to:

- identify major gaps in our knowledge of the flora and seek ways to fill those gaps expediently
- develop a detailed "Guide for Contributors" including working definitions of concepts such as escaped species, and formation and spelling of vernacular names
- pay close attention to making the Flora and database of maximal usefulness to users not trained in systematics or even in science
- provide to all associated groups information on the status of the project at all levels, e.g., who is committed to writing specific treatments, what the agendas of the Editorial Committee meetings are, etc.

IV. SUGGESTIONS FOR FURTHER INVESTIGATION

1. Fill in gaps of areas not adequately botanized.
2. Investigate possibility for line-item budgeting for FNA project.
3. Stay abreast of North American Biological Survey activities and attempt to coordinate with those.
4. Expand its role to include broad-based communication on floristic studies.

APPENDIX 11: WORKSHOP GOALS AND AGENDA

FLORISTICS FOR THE 21st CENTURY
4-7 MAY, 1988
Old Colony Inn, Alexandria, Virginia

Goals of the Workshop

1. Determine how floristic information is used now

2. Identify ways that use could be improved or increased

3. Determine what other uses could be made of floristic information, what other information should be gathered, how it might be presented

4. Determine what workforce, software, hardware, and management plan might be used to implement the above suggestions

5. Summarize the findings in a written report that can be used by institutions and funding agencies to help them understand the needs and importance of floristics

AGENDA

WEDNESDAY, 4 MAY: Presentations with question and answer periods. The purpose of this day is to help all participants understand what resources are available both for and from floristic studies and to understand the differences and similarities between various studies such as local, regional, and continental studies, checklists, and full Floras.

9:00-10:00 Welcoming remarks--Tod Stuessy, James Rodman, Nancy Morin; introductory remarks, overview of workshop agenda

10:00-10:20 Traditional floristic information and its uses--N. Morin
Sources of information and means of communication

10:20-10:40 Flora of North America Project--N. Morin et al.
1. Resources available to the project

10:40-11:00 Coffee Break

11:00-12:00 Flora of North America Project, continued
2. Current organization, schedule
3. Anticipated content of the published volumes
4. Anticipated content, structure, use, and availability of the database

12:00-1:30 Lunch Break

1:30-3:30 Current projects of various scopes (with discussion of computerization as appropriate)
1. Pennsylvania Atlas and Annotated Checklist--A. Rhoads
2. Illinois Natural History Survey--L. Nevling: Fish and Wildlife and mapping
3. Jepson's Manual--Hickman: user-oriented publication
4. Nature Conservancy--L. Morse: conservation
5. North Carolina Botanical Garden--J. Kartesz: Checklists and distributions

3:30-3:50 Coffee Break

3:50-5:00 Current Projects, continued
6. Specimen-based inventories used by zoologists--B. Chernoff
7. Flora of Southern Africa--G.E. Gibbs Russell: project organization
8. Flora Europaea--V. H. Heywood: European documentation system
9. Flora of Australia--A. George

5:00-5:30 Summary and Wrap-up

7:00-? Demonstrations of computer programs, informal session

THURSDAY, 5 MAY: General discussion and working group meetings. The purpose of this day is to explore how floristic information that is, or could be, gathered for Flora of North America can be used to fill specific needs and answer specific questions, and how it can be used to further stimulate research.

9:00-9:15 Introductory remarks, overview of day's activities

9:15-10:30 Traditional uses of floristic information using new technology

1. Evaluate floristic complements to characterize an area --G.E. Gibbs Russell
2. Distribution information for management--W. Brigham

10:30-10:40 Explanation of brainstorming technique and how small groups will be run.

10:40-11:00 Coffee Break

11:00-12:30	Working Group Meetings (see description after Lunch Break)
12:30-2:00	Lunch Break
2:00-3:30	Working group meetings. Groups will address the following aspects as they apply to their own work or to the purposes of their institutions, in each case asking: (1) What questions should be asked? (2) What information can be gathered as part of the FNA effort to answer the questions? (3) How can that information be presented most usefully?

Group 1. How is floristic information used in areas such as agriculture, forestry, and horticulture? How can this use be improved or increased?

Group 2. How is floristic information used to assess environmental conditions? How can this use be improved or increased?

Group 3. How is floristic information used to assess organisms' responses to changes in environmental conditions, e.g., expansion or contraction of ranges, vitality of populations, genetic changes?

Group 4. How is floristic information used in biological research outside of plant taxonomy? How can this use be improved or increased?

Group 5. How is information on characteristics and distributions of plants used to study phylogeny and biogeography?

3:30-3:50	Coffee Break
3:50-5:30	Summary of working groups and wrap-up: Representative from each working group presents conclusions of the group
7:00-9:00	Banquet

FRIDAY, 6 MAY: General discussion and working group meetings. The purpose of this day is to explore the practical implementation and ramifications of the needs and uses identified on the second day.

9:00-9:30	Opening remarks and overview of day's events; review of brainstorming technique
9:30-10:45	Brainstorming session to identify the possibilities
10:45-11:05	Coffee Break

11:05-12:30 Working group meetings

Group 6. Resources: What workforce is needed and available to gather information targeted by the previous day's working groups as desirable? How can the work force be mobilized and managed? What existing resources could be networked with the FNA database? Levels of data gathering include taxon-based, specimen-based, and living populations or individuals.

Group 7. Computer resources: What storage capacity will be needed for different levels of information? What software is available or needed for applications identified as desirable? What are considerations of access (ease, ownership, etc.)? How can we go about making information optimally transferrable and what software and hardware can we anticipate users having? How should networking be handled, and should permanent connections be established?

Group 8. How can and should the Flora of North America project interact and collaborate with projects of other scope, such as differing geographic levels, monographic works, etc.?

Group 9. What should the long-term management scheme of a project such as Flora of North America be in terms of its relationship to potential national/global biological surveys, in terms of its eventual composition (e.g., should it be a center for taxonomic plant information, for all plant-related information, or something intermediate), in terms of the need for perpetual maintenance and development of the data base?

Group 10. Additional topics identified in brainstorming session.

12:30-2:00 Lunch Break

2:00-3:40 Working groups, continued

3:40-4:00 Coffee Break

4:00-5:00 Reports from working groups

5:00-5:30 Wrap-up of day's activities

SATURDAY, 7 MAY: Summary of working group conclusions. Participants will spend day writing report from working group sessions. Afternoon meeting for summing up entire workshop--chaired by N. Morin.

9:00-10:00	Introductory remarks, explanation of format for group conclusions
10:00-10:45	Working groups
10:45-11:05	Coffee Break
11:05-12:30	Working groups, continued
12:30-2:00	Lunch Break
2:00-3:00	Working groups, continued
3:00-3:20	Coffee Break
3:20-5:00	Presentation of group conclusions
5:00-5:30	Summary of workshop: Review of workshop goals, accomplishments, and action to be taken

GROUP ASSIGNMENTS

* = leader, second day
' = recorder, second day
** = leader, third day
" = recorder, third day;

Alphabetical Listing

		day 2	day 3	write
Argus, George"	conservation	3	6	6
Barkley, Theodore'	floristics, regional	2	6	2
Beecher, Chris	databases: chemical	4	8	8
Bouchard, Andrè	phytogeography	5	10	
Brigham, Warren	databases: GIS	2	7	7
Brouillet, Luc'	research: phytogeography	3	6	3
Catling, Paul	applied: wetlands	4	8	4
Chernoff, B.	databases: collections			
Crosby, Marshall	databases: TROPICOS	5	9	9
Duncan, Thomas	databases	5	7	7
Duvick, Donald N.	applied: biotechnology	1	7	1
Estes, James**	floristics: grasses	5	9	9
Fay, John*	conservation	3	8	3
Flint, Harrison	applied: horticulture	3	10	3
George, Alex	floristics, continental	5	8	8
Gibbs Russell, G. E.	databases: floristics	4	6	4
Gibson, Arthur	research: plant physiology	3	6	6
Giles, Robert H.	databases: GIS, wildlife	1	7	7
Graham, R. L.	databases: GIS, envir. assess.	2	7	2

		day 2	day 3	write
Gunn, C. R.	applied: economic plants	1	7	1
Hall, David	applied: agricultural extension	1	10	1
Hartman, Ron"	floristics: plant taxonomy	5	8	8
Hayne, Don	applied: statistics	2	10	10
Haynes, Robert	databases: collections	5	7	5
Heywood, Vernon	conservation: floristics	3	9	9
Hickman, James	floristics, regional	4	8	4
Johnston, Marshall"	floristics, regional	4	7	7
Kartesz, John	databases: checklist	4	7	7
Kiger, Robert'	research: plant taxonomy	5	9	5
Kosztarab, Michael	res.: entom., nat'l. biol. surv.	4	9	9
Krugman, Stanley L.	applied: forest genetics	1	9	1
Larson, Gary	research: plant taxonomy	1	10	10
Leiser, Andrew	applied: cultivated plants	1	10	10
MacBryde, Bruce	conservation	3	10	10
Macrander, Michael**	databases	2	7	7
Magill, Robert	databases: TROPICOS	2	7	7
Miller, Norton	applied: state biol. survey	5	9	9
Morain, Stanley	databases: GIS	2	7	2
Morin, Nancy	floristics, continental			
Morse, Larry	databases: conservation	3	7	7
Murray, David"	databases: floristics	3	9	9
Nelson, C. J.	applied: agronomy	1	6	1
Nevling, Lorin	applied: state survey	5	8	5
Packer, John G.	research: plant taxonomy	1	10	10
Rhoads, Ann**	databases: state floristics	1	8	8
Richards, Jennifer**	research: plant morphology		4	6
Shetler, Stanwyn	databases: nat'l. biol. surv.	3	9	9
Shultz, Leila'	research: plant taxonomy	4	10	4
Simpson, Beryl*	research: ecology, popul. biol.	2	6	2
Spellenberg, Richard'	research: plant taxonomy	1	6	1
Stein, Bruce	conservation	2	9	9
Stuessy, Tod*	research: plant taxonomy	4	9	4
Taylor, Iain	conservation, plant genetics	4	10	4
Walter, Kerry	databases: living collections	3		
Waggoner, Gary**	database management: parks	2	10	10
Wagstaff, Jess	applied: epidemiology, toxicol.	1	8	8
Whetstone, R. David"	floristics: teaching	3	10	10
Wilken, Dieter*	floristics: plant taxonomy	5	6	5
Young, David*	applied: cultivated plants	1	8	1